

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 16:15:10 ; Search time 2064 Seconds
(without alignments)
2823.899 Million cell updates/sec

Title: US-09-993-808B-1

Perfect score: 1372

Sequence: 1 cccacggtccggacgaag.....agtgtcaactaaattcattc 1372

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1371.6	100.0	1372	7	ABV74603
2	644.2	47.0	723	3	AA95281
3	351	25.6	841	7	ABV74605
4	299	21.8	1242	6	ABK33958
5	234.6	17.1	639	3	AA95280
6	223.2	16.3	572	3	AA95280
7	223.2	16.3	572	3	AA95276
8	219	16.0	568	6	ABK33954
9	114.8	8.4	2015	2	AAZ24904
10	110.8	8.1	3313	5	AA9529102
11	110.8	8.1	3313	6	ABK33958
12	110.8	8.1	3313	9	ADC25236
13	110.2	8.0	1557	3	AA9529034
14	110.2	8.0	1557	5	AA9529030
15	109.6	8.0	701	4	AAH33453
16	108.8	7.9	1316	3	AA9529966
17	108.4	7.9	149	4	AA9529969
18	108.4	7.9	2116	3	AA9529969
19	108.4	7.9	2116	5	AA9529969
20	107.4	7.8	1837	3	AA9529969
21	107.4	7.8	2140	4	ABA06584
22	107.4	7.8	2140	4	ABA06584
23	107.4	7.8	2140	6	ABV83921

24	107.2	7.8	2322	7	ABX12257	Abx12257 cDNA enco
25	106.8	7.8	845	6	ABQ54253	Abq54253 Human ova
26	106.6	7.8	858	4	AAH33135	Aah33135 Human col
27	106.4	7.8	584	3	AA98302	Aac98302 Human col
28	106.4	7.8	603	5	ABA11175	Abal1175 Human ner
29	106.4	7.8	763	3	AA98303	Aac98303 Human col
30	106.4	7.8	824	2	AA98303	Aac98303 Human sec
31	105.6	7.7	2402	7	ABX12256	Abx12256 cDNA enco
32	105.4	7.7	418	3	AA98309	Aac98309 Human col
33	105.4	7.7	627	4	AA102583	Aal02583 Human rep
34	105.4	7.7	627	4	ABL97313	Ab197313 Human tes
35	105.4	7.7	1280	6	AAH34250	Aah34250 Human col
36	105.4	7.7	1280	6	ABQ54195	Abq54195 Human ova
37	105.4	7.7	1281	3	AA66412	Aac66412 Human sec
38	105.2	7.7	1928	6	ABK69104	Abk69104 DNA enco
39	105.2	7.7	2454	3	AA669584	Aac69584 Human sec
40	105.2	7.7	2454	7	ADC40257	Adc40257 Human sec
41	105.2	7.7	2454	9	ADC73781	Adc73781 Human sec
42	104.4	7.6	453	3	AA98320	Aac98320 Human col
43	104.4	7.6	540	3	AA98305	Aac98305 Human col
44	104.4	7.6	1270	3	AA959584	Aac95984 Human sec
45	104.4	7.6	2152	4	AA05156	Aad05156 Human sec

ALIGNMENTS

RESULT 1

ABV74603

ID ABV74603 standard; DNA; 1372 BP.

XX

AC ABV74603;

XX

DT 21-FEB-2003 (first entry)

XX

DE Maize CKI_B coding sequence.

XX

KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;

KW crop yield; root size; plant growth; tassel size; ear size;

KW male sterility; endoreduplication; gene; ds.

XX

OS Zea mays.

XX

PH Key Location/Qualifiers

FT CDS 134..904

FT /*tag= a

FT /product= "CKI_B"

XX

PN W0200281623-A2.

XX

PD 17-OCT-2002.

XX

PF 06-NOV-2001; 2001WO-US044038.

XX

PR 07-NOV-2000; 2000US-0246349P.

XX

PA (PION-) PIONEER HI-BRED INT INC.

PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX

Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

WPI; 2003-058511/05.

DR P-PSDB; ABB98757.

XX

Novel cyclin-dependent kinase polynucleotides and their encoded proteins, involved in cell cycle regulation, and useful for altering cell cycle protein content, cell cycle progression, cell number and composition of plants.

XX Claim 1; Page 64-65; 69pp; English.

PS

XX The present sequence is the coding sequence (I) for maize cyclin-

CC dependent kinase inhibitor (CKI), CKI_B. (I) is useful for modulating the

CC

activity of cyclin-dependent kinase (CDK) in a plant such as a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, oil-seed Brassica and millet. Modulating the activity of CDK, preferably modulating downward is useful for providing differential growth in a plant, especially a positive growth advantage and modulating CDK activity upward is useful for increasing crop yield, root size, plant growth, tassel size and/or ear size. Modulating CDK activity is also useful for conferring male sterility and for improving transformation frequencies by increasing the number of cells in cell division. CDK activity can also be modulated for modulating endoreduplication in the endosperm of corn, sorghum, wheat, rice, barley, and millet, where the promoter used is an endosperm-preferred promoter. The cell numbers are modulated in one or more tissues of a plant, comprising root, seed, tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves, stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm, pericarp, meristematic or leaf cells. (I) is also useful for identifying maize CKI interacting proteins, by adducting the nucleic acid sequence to a second nucleic acid sequence encoding a DNA-binding domain

Sequence 1372 BP; 292 A; 396 C; 442 G; 241 T; 0 U; 1 Other;

Query Match 100.0%; Score 1371.6; DB 7; Length 1372;
 Best Local Similarity 100.0%; Pred. No. 1.5e-270; Indels 0; Gaps 0;
 Matches 1372; Conservative 0; Mismatches 0;

Qy 1 CCCACGGCTCCGACGCAAGCGCTGCGAGGAGAGCGCGCGAGGCGTGTGGCCCTGT 60
 Db 1 CCCACGGCTCCGACGCAAGCGCTGCGAGGAGAGCGCGCGAGGCGTGTGGCCCTGT 60

Qy 61 GGGAGAGGMAAAG 120
 Db 61 GGGAGAGGMAAAG 120

Qy 121 GGGCTTCGCTCAGATGGGAGAGTACATCGCAAGTGCAGAGGCGCGAGCGCGGAGGT 180
 Db 121 GGGCTTCGCTCAGATGGGAGAGTACATCGCAAGTGCAGAGGCGCGAGCGCGGAGGT 180

Qy 181 GCGCGCGCTCAGGTTACGAGGTCGTGCGCGTCCGAGAGGTCAGGTCGCGCGCGG 240
 Db 181 GCGCGCGCTCAGGTTACGAGGTCGTGCGCGTCCGAGAGGTCAGGTCGCGCGCGG 240

Qy 241 GACCGCGGTGTCGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 Db 241 GACCGCGGTGTCGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

Qy 301 TGCGCGCTGAGCGCTGGTGGGACGCGGAGAGTCTACATCACTGCTAGCGGCAT 360
 Db 301 TGCGCGCTGAGCGCTGGTGGGACGCGGAGAGTCTACATCACTGCTAGCGGCAT 360

Qy 361 GCTGTTTCATGGCAACCGCTCAGCGCGAGCGCTGGTTCCTGGTTCAGCGCGGTTGA 420
 Db 361 GCTGTTTCATGGCAACCGCTCAGCGCGAGCGCTGGTTCCTGGTTCAGCGCGGTTGA 420

Qy 421 GGTGCTGATGGCGCTCAG 480
 Db 421 GGTGCTGATGGCGCTCAG 480

Qy 481 CTCACGACGCGCTCGTTCGTTGAACTTGGGGTGGGGGTGAGCGCGGAGAGAGAGAGAG 540
 Db 481 CTCACGACGCGCTCGTTCGTTGAACTTGGGGTGGGGGTGAGCGCGGAGAGAGAGAGAG 540

Qy 541 CGCTCTTCAGAGCTGAG 600
 Db 541 CGCTCTTCAGAGCTGAG 600

Qy 601 GAGCAACTCCGGAG 660
 Db 601 GAGCAACTCCGGAG 660

Qy 661 GCACGCGAGCTCAGCGATCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 720
 Db 661 GCACGCGAGCTCAGCGATCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 720

Qy 721 ACCGCGCGCAACGCCGCTCGGAGCTGATCGTCCGCCGACACACAGATCCAGAGTT 780
 Db 721 ACCGCGCGCAACGCCGCTCGGAGCTGATCGTCCGCCGACACACAGATCCAGAGTT 780

Qy 781 CTTCCGCGCGCCGAGGCGGCCAGGCGCCAGGCGCTTCTTCCAGTACAACTTCGACTT 840
 Db 781 CTTCCGCGCGCCGAGGCGGCCAGGCGCCAGGCGCTTCTTCCAGTACAACTTCGACTT 840

Qy 841 CGTCCGCGCGCTGCGCTCGACGCGCGCGCGCTTCGAGTGGCGCGCGTGGTCAAGAT 900
 Db 841 CGTCCGCGCGCTGCGCTCGACGCGCGCGCGCTTCGAGTGGCGCGCGTGGTCAAGAT 900

Qy 901 CTGAAGCGAGCGCTCGCTCGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 960
 Db 901 CTGAAGCGAGCGCTCGCTCGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 960

Qy 961 CCCCCCCCCCAACAACATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 Db 961 CCCCCCCCCCAACAACATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020

Qy 1021 GCCTAAGCTAACACACACCATTCATCTCGTCCAAATGATGCTTCTTCTCTGGAGC 1080
 Db 1021 GCCTAAGCTAACACACACCATTCATCTCGTCCAAATGATGCTTCTTCTCTGGAGC 1080

Qy 1081 TAGCAGGAGCGTGTATTTATTTAGTACTACTTTTACTTATTTACAGAGTTATCTTGACCC 1140
 Db 1081 TAGCAGGAGCGTGTATTTATTTAGTACTACTTTTACTTATTTACAGAGTTATCTTGACCC 1140

Qy 1141 GATAGATCAATPCGCTTACTGTGTAATTTCTTCATCATCTCTTAGATGGAGTTTAAATC 1200
 Db 1141 GATAGATCAATPCGCTTACTGTGTAATTTCTCTCATCATCTCTTAGATGGAGTTTAAATC 1200

Qy 1201 GTCTTATTTATTTACTGTACAGAGCTTGTGGCTTGGCTTGGCAAGAGATCTGGTTCT 1260
 Db 1201 GTCTTATTTATTTACTGTACAGAGCTTGTGGCTTGGCTTGGCAAGAGATCTGGTTCT 1260

Qy 1261 CAAAAAAG 1320
 Db 1261 CAAAAAAG 1320

Qy 1321 TAGCGTGCATGCGAGCTCATAGCTCTTCTATAGTGTACCTAAATTCATTTC 1372
 Db 1321 TAGCGTGCATGCGAGCTCATAGCTCTTCTATAGTGTACCTAAATTCATTTC 1372

RESULT 2

AAA95281 ID AAA95281 standard; cdna; 723 BP.
 AC AAA95281;
 DT 17-JAN-2001 (first entry)

XX
 AC AAA95281;
 DT 17-JAN-2001 (first entry)

XX
 DE Corn cyclin-dependent kinase inhibitor coding sequence #3.

XX
 KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 cell growth; herbicide; ss.

XX
 OS Zea mays.

XX
 FH Key Location/Qualifiers
 CDS 3..380
 FT FT /*tag= a
 FT FT /product= "CDKI"
 FT FT /Partial

XX
 FN WO200060087-A2.

XX
 PD 12-OCT-2000.

XX
 PF 06-APR-2000; 2000WO-US009106.

XX
 PR 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA Klein TM, Weng Z, Cahoon RE;
PI WPI; 2000-679375/66.
XX P-PSDB; AAB26250.
XX
XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
XX Claim 2; Page 43-44; 58pp; English.
XX
XX The present sequence is the coding sequence for the corn cyclin-dependent
CC kinase inhibitor (CKI). It was isolated by searching a contig comprising
CC cDNA from corn roots and ear leaf sheath for sequences similar to those
CC encoding the CKI from *Chenopodium rubrum*, *Caenorhabditis elegans* and
CC Arabidopsis thaliana. CKI is involved in the cell cycle, and may promote
CC or inhibit cell division and growth. The coding sequence and the protein
CC it encodes are useful in the production of transgenic plants which
CC produce increased or decreased amounts of the CKI protein, in the
CC identification of herbicides, in genetic and physical mapping and in the
CC isolation of the CKI gene in other organisms
XX
SQ Sequence 723 BP; 160 A; 202 C; 210 G; 151 T; 0 U; 0 Other;

Query Match 47.0%; Score 644.2; DB 3; Length 723;
Best Local Similarity 95.9%; Pred. No. 4.8e-122;
Matches 707; Conservative 1; Mismatches 14; Indels 15; Gaps 4;

QY 525 GCGGGAGCCACACTGCGCGTCTTACGACGCTGCGAGGCTGGCGGGATCAGCTCTCTGG 584
DB 1 GCGGGAGCCACACTGCGCGTCTTACGACGCTGCGAGGCTGGCGGGATCAGCTCTCTGG 60

QY 585 TGGATGCTCTCGCGGGGAGCACTCCGGAGCGGCCAGACCGCGAGCGAGAGACGA 644
DB 61 TGGTGGATGCTCTCGCGGAGCACTCCGGAGCGGCCAGACCGCGAGCGAGAGACGA 120

QY 645 CGCATCGAGCGGGGCGAGCGGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACCA 704
DB 121 CGCATCGAGCGGGGCGAGCGGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACCA 180

QY 705 AGACTGCCGCTCGTACCGGCGCAACCGCGCTGCGAGCTCATCTGTCGCGCCAGCAC 764
DB 181 AGACTGCCGCTCGTACCGGCGCAACCGCGCTGCGAGCTCATCTGTCGCGCCAGCAC 240

QY 765 ACGAGATCCAGGAGTTCTTGCCGCCCGCGAGCGGCCAGCGCTTTCCTTCCA 824
DB 241 ACGAGATCCAGGAGTTCTTGCCGCCCGCGAGCGGCCAGCGCTTTCCTTCCA 300

QY 825 AGTAACTTCGACTTCTGCGGGCGGTGCCCTCGAGCGCGCGCGGTTCGAGTGG 884
DB 301 AGTAACTTCGACTTCTGCGGGCGGTGCCCTCGAGCGCGCGCGGTTCGAGTGG 360

QY 885 CGCGGTGGTTCAGCATCTGAAGCGAGCGTGGCTCGGTGCAAGGTGAGTGAAGAAGA 944
DB 361 CGCGGTGGTTCAGCATCTGAAGCGAGCGTGGCTCGGTGCAAGGTGAGTGAAGAAGA 420

QY 945 AAAGATCCCCCCCCCCCCCCCCCAACAACATACGAGGAGAGAAACCAACAAT 1004
DB 421 AAAGATG-----CCCCCAACAACAATAACGAGAGAG-AAAACCGAACAAT 467

QY 1005 TAACGAGTTTATAGCTTAAGCTTAACCAACGATTCATCTCGTCAATGATGCT 1064
DB 468 TAACGAGTTTATAGCTTAAGCTTAACCAACGATTCATCTCGTCAATGATGCT 527

QY 1065 TGCTTTCTCTGGAGCTAGCAGGAGCGTAGTATTATTAGTACTACTTACTTATTTCAG 1124
DB 528 TGCTTTCTCTGGAGCTAGCAGGAGCGTAGTATTATTAGTACTACTTACTTATTTCAG 587

QY 1125 AGGTTATCTTACCCCGATAGATCAATCCGTTACTGTGTAATTTCTCTCATGATCTCT 1184
DB 588 AGGTTATCTTACCCCGATAGATCAATCCGTTACTGTGTAATTTCTCTCATGATCTCT 647

QY 1185 TAGATGGAGTTTAATCGTCTTAATTTTACTCTACAGCAGCTTGTTGGCTTGCRAAGA 1244
DB 648 TAGATGGAGTTTAATCGTCTTAATTTTACTCTACAGCAGCTTGC-TGGCTTGCRAAGA 706

QY 1245 AAGATCTGGTTTGTCTC 1261
DB 707 AAG-TCTGGTTTGTCTC 722

RESULT 3
ABV74605
ID ABV74605 standard; DNA; 841 BP.
XX
XX AC ABV74605;
XX DT 21-FEB-2003 (first entry)
XX DE Maize CKI_D coding sequence.
XX KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_D; plant; corn;
XX crop yield; root size; plant growth; tassal size; ear size;
XX male sterility; endoreduplication; gene; ds.
XX OS Zea mays.
XX FH Key Location/Qualifiers
FT CDS 159..839
FT /tag= a
FT /partial
FT /product= "CKI_D"
FT /note= "No stop codon given"
XX DN WO200281623-A2.
XX PD 17-OCT-2002.
XX PF 06-NOV-2001; 2001WO-US044038.
XX PR 07-NOV-2000; 2000US-0246349P.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
XX DR P-PSDB; ABB98759.
XX
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX Claim 1; Page 68-69; 69pp; English.
XX
XX The present sequence is the coding sequence (1) for maize cyclin-
CC dependent kinase inhibitor (CKI), CKI_D. (1) is useful for modulating the
CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
CC preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassal size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassal, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,

XX 01-NOV-2001 (first entry)
 DT Corn Cyclin dependent kinase inhibitor (CDKI) clone csiln.pk0050.e6.
 DE Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; corn;
 KW plant growth inhibitor; ds.
 XX Zea mays.
 OS
 FH Key Location/Qualifiers
 FT CDS 208..369
 FT /*tag= a
 FT /product= "CDKI fragment"
 FT /partial
 FT /note= "No start or stop codon given"
 XX
 XX
 PN WO200060087-A2.
 XX
 PD 12-OCT-2000.
 XX
 PF 06-APR-2000; 2000WO-US009106.
 XX
 PR 07-APR-1999; 99US-0128192P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA Klein TM, Weng Z, Cahoon RE;
 PI WPI; 2000-679375/66.
 XX P-PSDB; AAP01940.
 DR
 XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.
 XX
 PS Claim 2; Page 39-40; 58pp; English.
 XX
 CC The invention describes a novel isolated polynucleotide comprising a
 CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
 CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
 CC of cell division, growth and death. The nucleotide sequences can be used
 CC in a vector to transform a host cell to produce the CDKI polypeptide.
 CC They can also be used in methods for selecting and obtaining a nucleic
 CC acid sequence that encodes CDKI or affects the level of CDKI expression.
 CC The encoded protein can be used in a method for evaluating a compound for
 CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
 CC as herbicides. They can also be used to inhibit plant growth. The
 CC polynucleotide sequences can be used in gene mapping and as genetic
 CC markers. The sequence encodes the corn CDKI clone csiln.pk0050.e6 as
 CC described in the method of the invention
 XX
 SQ Sequence 572 BP; 140 A; 156 C; 176 G; 93 T; 0 U; 7 Other;
 Query Match 16.3%; Score 223.2; DB 3; Length 572;
 Best Local Similarity 78.5%; Pred. No. 4.3e-36;
 Matches 335; Conservative 0; Mismatches 74; Indels 18; Gaps 5;
 537 CCTGCGCTCCTACGACGCTGACAGGCTGGCGGGATCAGCTCCTGGTGTGCTCGG 596
 2 CCTGCGCTCGACGCTGGCTGGAGTTGACGGGATCAGTCCGCGATGCTGCA 61
 597 CGGCGAGCACTCCGGAGCGGCCAGACCGCGAGCGGAGACGACGCCA---TCGA 653
 62 CCGGAGCACTCCGGAGCGCTCCCGGACCGCGAGGAGAGACGACGCTCGTA 121
 654 GCCGGGCGCA---CGGCGAGTTCAGGATCGAGTCTGGCGG---GGCACAGA 707
 122 GCCGGGCGCACGGCGCGGAGTTCAGGATCGAGTCTGGTGGGGCGGAGAGA 181
 708 CTGGC-----CCGTCGTACCGCGGCAACCGCGCTGCGGAGTGTGTCGCCAG 761
 182 CTGGCTGCTGCTGCTGCGCGCGGACACAAATCGTTGGGAGTGTGTCGCCAG 241

QY 762 CACACGAGATCCAGAGTTCTTTCGCGCGCGCGGCGCCAGCGGCGCCAGCGCTTTCCTT 821
 Db 242 CACAGGAGATCCAGGAAATCTTCGCGCGCGCGGCGGCGCCATGCCAAACGCTTTCCTT 301
 QY 822 CCAAGTACAACTTCGACTTCGCTCCGCGCGCTGCCCTTCGACCGCGCGCTTCGAGT 881
 Db 302 CCAAGTACAACTTCGACTTCGCTCCGCGCGCTGCCCTTCGACCG---CGGCGGCTTCGAGT 358
 QY 882 GGGCGCGGTGGTTCAGCATCTGAGCGAGCGTGCCTCCGCTCAAGTCAAGCTAGAAAG 941
 Db 359 GGAGCGCCANGGTTCAGCATCTGAGCGAGCTCGGTGCAAGGTGAACCTACTACTAGTAG 418
 QY 942 AGAAAAG 948
 Db 419 AAATGAG 425
 RESULT 7
 AAA95276
 ID AAA95276 standard; cDNA; 572 BP.
 AC AAA95276;
 XX
 DT 17-JAN-2001 (first entry)
 XX
 DE Corn cyclin-dependent kinase inhibitor coding sequence #1.
 XX
 KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 KW cell growth; herbicide; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT CDS 208..366
 FT /*tag= a
 FT /product= "CDKI"
 FT /partial
 FT
 FT
 FT
 XX WO200060087-A2.
 PN
 PD 12-OCT-2000.
 XX
 PF 06-APR-2000; 2000WO-US009106.
 XX
 PR 07-APR-1999; 99US-0128192P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA Klein TM, Weng Z, Cahoon RE;
 XX WPI; 2000-679375/66.
 DR P-PSDB; AAB26245.
 XX
 XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.
 XX
 PS Claim 2; Page 39-40; 58pp; English.
 XX
 CC The present sequence is the coding sequence for the corn cyclin-dependent
 CC kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA
 CC library for sequences similar to those encoding the CDKI from Chenopodium
 CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
 CC in the cell cycle, and may promote or inhibit cell division and growth.
 CC The coding sequence and the protein it encodes are useful in the
 CC production of transgenic plants which produce increased or decreased
 CC amounts of the CDKI protein, in the identification of herbicides, in
 CC genetic and physical mapping and in the isolation of the CDKI gene in
 CC other organisms
 XX
 SQ Sequence 572 BP; 140 A; 156 C; 176 G; 93 T; 0 U; 7 Other;
 Query Match 16.3%; Score 223.2; DB 3; Length 572;
 Best Local Similarity 78.5%; Pred. No. 4.3e-36;

Matches 335; Conservative 0; Mismatches 74; Indels 18; Gaps 5;

QY 537 CTTGCGCTCCTACGAGCTGAGAGCTGGCGGGGATACGTCCTGGTGGATGCTCGG 596
 DB 2 CTTGCGCTCCTACGAGCTGAGAGCTGGCGGGGATACGTCCTGGTGGATGCTCGG 61
 QY 597 GCGCGAGCAACTCCGGAGCGGCCAGACCGCGAGAGCGGAGAGAGAGCGCCA---TCGA 653
 DB 62 CCGCGAGCAACTCCGGAGCGGCCAGACCGCGAGAGAGAGAGAGAGAGAGAGAGAG 121
 QY 654 GCGCGCGCA---CGCGAGCTCAGCGATCTGAGTGGAGTCTGGCGG---GGCAAGA 707
 DB 122 GCGCGCGCAAGCGCGAGCTCAGCGATCTGAGTGGAGTCTGGTGGCGCGCAAGA 181
 QY 708 CTGGC-----CGTGTCTACCGCGGCAACCGCGGCTGGAGCTGATGTCGCCCG 761
 DB 182 CTGGCTGCTGCTGCTGCGCGGCAACAACATCGGTTGCGAGCTGATGTCGCCCG 241
 QY 762 CACACGAGATCCAGGAGTCTTCCGCGCGCCGAGCGCGGCCAGGCGCAAGCGCTTTGCTT 821
 DB 242 CACAGAGATCCAGGAGTCTTCCGCGCGCCGAGCGGCCAGGCGCAAGCGCTTTGCTT 301
 QY 822 CCAAGTACAACTTCGACTTGTCTCGCGGGTGCCCTCGACGCGCGCGGCTTCGAGT 881
 DB 302 CCAAGTACAACTTCGACTTGTCTCGCGGGTGCCCTCGACGCGCGCGGCTTCGAGT 358
 QY 882 GCGCGCGGCTGTCAGCATCTGAAGAGCGTGCCTCGGTCGAAGGTGAAGTAGAAG 941
 DB 359 GGAAGCGGAGGTCAGCATCTGAAGAGCGTGCCTCGGTCGAAGGTGAAGTAGAAG 418
 QY 942 AGAAAG 948
 DB 419 AATGAG 425

RESULT 8
 ABK93954
 ID ABK93954 standard; DNA; 568 BP.

AC ABK93954;
 XX
 XX 07-AUG-2003 (revised)
 DT 27-AUG-2002 (first entry)
 XX
 XX Corn zmlCK1 EST DNA sequence.
 DE
 XX Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.

XX Zea mays.

XX WO200228893-A2.

XX 11-APR-2002.

XX 29-JUN-2001; 2001WO-IB001492.

XX 14-JUL-2000; 2000US-0218471P.

XX 13-OCT-2000; 2000US-0241219P.

XX (CROP-) CROPDESIGN NV.

XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 PI Hatzfeld Y;

XX WPI; 2002-471311/50.
 DR P-PSDB; ABG65674.

XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.

XX Claim 41; Page 125; 141pp; English.

XX This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a cDNA molecule
 CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
 CC invention. (Updated on 07-AUG-2003 to correct OS field.)

XX Sequence 568 BP; 139 A; 149 C; 177 G; 102 T; 0 U; 1 Other;

Query Match 16.0%; Score 219; DB 6; Length 568;

Best Local Similarity 84.6%; Pred. No. 3.1e-35;

Matches 335; Conservative 0; Mismatches 36; Indels 25; Gaps 7;

QY 563 GCTGGCGGGGATCACGCTCCTGGTGGATGCTCGCGCGGAGCAACTCCGGAGCGGCCCA 622
 DB 2 GGTACGGGATCACGTCC---CGATGTCGTCTCCGCGAGCAACTCGGGAGCGTCCCG 58
 QY 623 GACCGCGAGAGCGGAG 676
 DB 59 GACCGCGAG 118
 QY 677 GATCTGAGTGGATCTGGCGG---GGCACAAGACTGGC-----CGTCTGCTACCGCG 727
 DB 119 GATCTGAGTGGATCTGGTGGCGGCGAGAGACTGGTCTGCTCGTCTCGCGGAGACA 178
 QY 728 GCAACGCGGCTGCGGAGCTGATCGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 787
 DB 179 ACAACATCGGCTCGGAGAGTATCGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 238
 QY 788 GCGCGCGAGCGCGCCAGAGCGCTTTGCTTCCAAGTACAACTTCGACTTCGTCGCG 847
 DB 239 GCGCGCGAGCGCGCCCACTGCCAAACGCTTTGCTTCCAAGTACAACTTCGACTTCGTCGCG 298
 QY 848 GCGTGTCCCTCGACGCGCGCGCGGCTTCGAGTGGCGCGCGGCTTCAGCATCTGAGC 907
 DB 299 GCGTGTCCCTCGACGCGCGCGCGGCTTCGAGTGGAGCGCGCGGCTTCAGCATCTGAGC 355
 QY 908 GAGCGTGGTTCGAGTGCAGGTGAAGCTAGAAAG 943
 DB 356 GAGCGTG-----CGTGCAGGTGAAGCTACTACTAG 387

RESULT 9
 AAZ24904

ID AAZ24904 standard; DNA; 2015 BP.

XX AAZ24904;

XX 02-DEC-1999 (first entry)

XX Human secreted protein gene 94 clone HLYCH68.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; retinosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX OS

XX WO9947540-A1.

XX 23-SEP-1999.

```

PF 18-MAR-1999; 99WO-US005804.
XX
XX 19-MAR-1998; 98US-0078563P.
PR 19-MAR-1998; 98US-0078563P.
PR 19-MAR-1998; 98US-0078563P.
PR 19-MAR-1998; 98US-0078573P.
PR 19-MAR-1998; 98US-0078574P.
PR 19-MAR-1998; 98US-0078576P.
PR 19-MAR-1998; 98US-0078577P.
PR 19-MAR-1998; 98US-0078578P.
PR 19-MAR-1998; 98US-0078579P.
PR 19-MAR-1998; 98US-0078581P.
PR 01-APR-1998; 98US-00803112P.
PR 01-APR-1998; 98US-00803113P.
PR 01-APR-1998; 98US-00803114P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
PI Wei Y, Redress CA, Duan RD, Kyaw H, Ebner R, Lafleur DW, Olsen HS;
PI Shi Y, Moore PA;
XX
XX WPI; 1999-562050/47.
DR P-PSDB; AAY41401.
XX
XX New isolated human genes, useful for diagnosis and treatment of e.g.
PT cancers, neurological disorders, immune diseases, inflammation or blood
PT disorders.
XX
XX Claim 1; Page 354-355; 484pp; English.
XX
XX This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number, and the clone it is derived from, are
CC detailed in the descriptor line. The gene can be used to generate fusion
CC proteins by linking to the gene to a human immunoglobulin Fc portion
CC (e.g. AAZ24802) for increasing the stability of the fused protein as
CC compared to the human protein only. The invention relates to 95 novel
CC genes and their fragments (nucleic acid sequences: AAZ24811-224907; amino
CC acid sequences AAY41308-Y41404) which are useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 95 polynucleotides, based on which tissues they are most highly expressed
CC in (see AAZ24811 for described uses)
XX
XX Sequence 2015 BP; 536 A; 500 C; 508 G; 482 T; 0 U; 9 Other;
SQ
Query Match 8.4%; Score 114.8; DB 2; Length 2015;
Best Local Similarity 86.8%; Pred. No. 7.9e-14;
Matches 125; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
Qy 1228 TGGTTGGCTTCAAGAAAGATCTGGTTTGTCTCAAAAAAATAAAAAAATAAAAAA 1287
Db 1514 TGTGTAGGATTAATAAACCACTTGTATGGAAAAAATAAAAAAATAAAAAA 1573
Qy 1288 AAGCGCGCCGCTTAGAGGATCCAAAGTTACGTACGCGTCGACGTCATAGCTCT 1347
Db 1574 AAGCGCGCCGCTTAGAGGATCCAAAGTTACGTACGCGTCGACGTCATAGCTCT 1633
Qy 1348 TCTATAGTGTCACCTAATTCATT 1371
Db 1634 TCTATAGTGKACCTAATTCATT 1657
RESULT 10
AAS29102
ID AAS29102 standard; cDNA; 3313 BP.
XX
AC AAS29102;
XX
XX 21-NOV-2001 (first entry)
XX
XX cDNA encoding for human DNA-binding protein #73.

```

```

XX Human; DNA-binding protein; histone; chromo domain protein;
KW chromatin organisation modifier; Y-box binding protein; DNA organisation;
KW gene transcription; malignant disease; autoimmune disorder;
KW rheumatic disease; genetic abnormality; infectious disease;
KW neurological disorder; gene therapy; immunomodulatory; anti-HIV;
KW anti rheumatic; anti microbial; cytostatic; ss.
XX
XX Homo sapiens.
XX
XX WO200155162-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001305.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224516P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225286P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.

```

```
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465557/50.
XX P-PSDB; AAU18226.
XX Nucleic acid molecules encoding human secreted chromosomal binding
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers.
XX Claim 4; SEQ ID NO 83; 561pp; English.
XX The present invention relates to the isolation of novel DNA-binding
CC proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding for
CC these proteins. DNA-binding proteins such as histones, chromo (chromatin
CC organisation modifier) domain proteins, and Y-box binding proteins may
CC contribute to diseases resulting from aberrant DNA organisation and/or
CC gene transcription. The sequences of the invention are useful in
CC screening assays to identify antagonists and/or agonists that may enhance
CC or block activities mediated by DNA-binding proteins. Blockers of DNA-
CC binding proteins may be useful in treating disorders such as malignant
CC diseases (e.g. cancer), autoimmune disorders (e.g. diabetes mellitus),
CC rheumatic diseases (e.g. rheumatoid arthritis), genetic abnormalities
CC (e.g. cystic fibrosis), infectious diseases (e.g. HIV) and neurological
CC disorders (e.g. Alzheimer's disease). The polynucleotide sequences of the
CC invention may also be used in gene therapy. AAS29030-AAS29157 represent
CC cDNA sequences encoding for novel DNA-binding proteins. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3313 BP; 1042 A; 693 C; 582 G; 991 T; 0 U; 5 Other;

Query Match 8.1%; Score 110.8; DB 5; Length 3313;
Best Local Similarity 89.4%; Pred. No. 5.8e-13;
Matches 118; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1240 AAGAAGATCTGGTTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGC 1299
Db 3107 AATAAACATTTGGATTTTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGC 3166

QY 1300 TCTAGAGGATCCAGCTTAGTACGCGTGCATGCGAGTCTCTTAGTGTCA 1359
Db 3167 TCTANAGGATCCAGCTTAGTACGCGTGCATGCGAGTCTCTTAGTGTCA 3226

QY 1360 CCTAAATTCATT 1371
Db 3227 CCTAAATTCAT 3238

RESULT 11
ABS68242
ID ABS68242 standard; cDNA; 3313 BP.
XX ABS68242;
AC ABS68242;
XX 18-NOV-2002 (first entry)
XX cDNA encoding human DNA-binding protein #73.
XX Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;
KW severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;
KW diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;
KW graft-versus-host disease; blood-related disorder; atherosclerosis;
KW hyperproliferative disorder; cancer; renal disorder; arrhythmia;
```


XX 13-MAR-2003.
XX 07-MAR-2002; 2002US-00091483.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218280P.
PR 26-JUL-2000; 2000US-0220363P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 13-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764846.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

PI

QY	1300	TCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCGACGTATAGCTCTTCTATAGTGCA	1359
Db	495	TCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCGACGTATAGCTCTTCTATAGTGCA	554
QY	1360	CCTAAATTCATT	1371
Db	555	CCTAAATTCAT	566

Search completed: October 1, 2004, 17:46:24
Job time : 2068 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 16:15:10 ; Search time 19992 Seconds
(without alignments)
2974.522 Million cell updates/sec

Title: US-09-993-808B-1

Perfect score: 1372

Sequence: 1 ccacgcgtccgagcgaag.....agtgccacctaattcattc 1372

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	301	21.9	1266	8	AK103084 Oryza sat
2	299	21.8	1242	6	AX406686 Sequence
3	219	16.0	568	6	AX406682 Sequence
c	158.4	11.5	135670	8	AP005002 Oryza sat
5	114.8	8.4	2015	6	BD136416 95 human
6	108.4	7.9	2658	5	DRES05990 Danio rer
7	107.2	7.8	2322	6	AX670965 Sequence
8	106.4	7.8	824	6	BD072967 70 human
9	105.6	7.7	2402	6	AX670963 Sequence
10	105.2	7.7	2454	6	BD276004 62 Human
11	104.4	7.6	1270	6	BD275352 49 Human
12	101	7.4	2657	6	BD275987 62 Human
13	97.2	7.1	1450	6	BD247539 Methods f
14	97	7.1	2486	6	AX301825 Sequence
15	95.6	7.0	1581	6	AR066494 Sequence
16	95.6	7.0	1581	6	I60018 Sequence 1
17	95	6.9	1255	6	AR137827 Sequence
18	95	6.9	1255	6	AR170186 Sequence
19	94.4	6.9	748	5	AU301081 Cyprinus
20	93.8	6.8	1495	6	BD247537 Methods f
21	93.4	6.8	493	6	AX406678 Sequence
22	93.2	6.8	1984	6	BD230989 49 human
23	93.2	6.8	1984	6	AR337284 Sequence
24	92.6	6.7	1464	6	BD270355 50 human
25	91.6	6.7	1665	6	AR260531 Sequence
26	91.6	6.7	1665	6	AX058884 Sequence
27	91	6.6	944	6	AR184178 Sequence
28	91	6.6	944	6	BD192707 123 human
29	88.8	6.5	2581	6	BD130531 Ligand re
30	87.6	6.4	2761	6	AX100254 Sequence
31	87	6.3	681	6	AX406731 Sequence
32	85.4	6.2	555	5	OLA238010 Oryzias 1
c	84.6	6.2	125020	9	AF429315 Homo sapi
33	84	6.1	696	6	AR184125 Sequence
34	84	6.1	696	6	BD192654 123 human
35	84	6.1	696	6	BD230968 49 human
36	82.4	6.0	831	6	BD230968 49 human
37	82.4	6.0	831	6	AR337263 Sequence
c	82	6.0	1330	6	AR137825 Sequence
38	82	6.0	1330	6	AR170184 Sequence
c	82	6.0	1653	6	BD249962 50 human
40	81.4	5.9	1813	6	BD233720 31 human
41	81.4	5.9	1813	6	AX780203 Sequence
42	80.8	5.9	3325	6	AU301063 Cyprinus
43	80.6	5.9	425	5	BD211760 Proteins
44	80.6	5.9	2819	6	AX592621 Sequence
45	80.4	5.9	1668	6	AX592621 Sequence

ALIGNMENTS

RESULT 1

AK103084

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Pred. No. is the number of results predicted by chance to have a

AK103084 1266 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J03118G18, full
insert sequence.

AK103084.1 GI:32988293

FLI CDNA; CAP trapper.

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaeae; Oryza.

1

The Rice Full-Length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-Length cDNA Project Team:
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group, Ohtsuki, K., Murauchi, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Maeda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
 Sato, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice
 Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2 (bases 1 to 1266)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
 Namiki, T., Narikawa, R., Niikura, J., Nishii, K., Nomura, K.,
 Numasaka, R., Ohneda, E., Ohno, M., Ohtsuki, K., Ooka, M., Ooka, H.,
 Osato, N., Ota, Y., Ohtsuki, K., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
 Yoshimura, A.

Direct Submission
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica
 rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
 Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
 Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
 Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishii, K., Nomura, K., Numasaka, R., Ohno, M., Osato, N.,
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,

Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
 Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
 1. 1266
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J0333118G18"

ORIGIN

Query Match 21.9%; Score 301; DB 8; Length 1266;
 Best Local Similarity 67.5%; Pred. No. 3e-41;
 Matches 585; Conservative 0; Mismatches 220; Indels 62; Gaps 9;

QY	113	AGGGCGCGCGGTCGTCAGATGGGAGTACATGCGCAAGTGCAGGGGGCCGCGAGGC	172
DB	40	AGGGCGCGCGCACGCGCGCAGATGGCAAGTACATGAGAAAGTTTCAGGGGGCCGCGAGGC	99
QY	173	GCGGAGGTCGCGCGCGTCGAGGTTACGAGGTCGTCGCGCGTCGCGACGAGGTCAGGTC	232
DB	100	GAGGATTTGGCGCGCATGAGGTCACGAGGTTGGCGTCGCGACGAGGTCGAGGTCG	159
QY	233	GCGGCGCGCACCGCGCG-----GTGTCGCGAAAGTCCGCCCGAGGAGG	274
DB	160	GCAGCGCGCGCGCGCGCGACGACGAGGTCGAGGCGCGCTCGCGCGCTCCACAGG	219
QY	275	AAGAGGCGCGCGCGCGCGGAGCCTGCTCCGCCGTCGAGGTCGTCGCGGAGCGCGGAGGC	334
DB	220	AGGAGGAGGCGGCTGCTCCGACGCGCGTCGTCGCGGACTCTCGCGCGTCGAGCGCGGAGC	279
QY	335	TGCTATATCCACCTGCGTAGCGGATGCTGTTTCATGCGCACCGCTTCAGCGCGAGCGTCG	394
DB	280	TGCTATCTCCAGCTGAGGAGCGGATGCTGTTTCATGCGCGCGCGC---GAGCGCGCGCG	336
QY	395	GTGACTCGGTTCCGACCGCGCGTGGAGGCTCTGATGCGGTCGTCGAGGACAGCGGCGCG	454
DB	337	GCGCGAGGCGTCCGCTGTTAGCGGCGCGGCTTCGCGGACGAGCGCGCGCGCAT	396
QY	455	GCGCTCGCGCGCGCGCTCTCGCGTCTCGAGCGCGGCTGCTCGTGAACCTTGCGCTG	514
DB	397	GCG---GCGGCTGCGCTCTCGCGTCTCGAGCGCGGCTGTC-----CGTGACGCG	447
QY	515	GCGGCTCAGCGCGCGGAGCACACCTCGCGCTCTACGACGCTGTCAGAGGTCGCGCGGAT	574
DB	448	GCGGCTCAGGACAGGAGGCTCGCGTGGCGCTC---CGAGCTCGCGGAGGCGAGCGAG	504
QY	575	CACGTCCTGGTGGATGTCGCGCGCGGAGCAATCTCGGAGGCGGCGCCAGACCGCGAGAGG	634
DB	505	CATGTCGCGAGG-----GCTCCGCGAGCGACTCGCGGAGCGCGCGTGCACCGCGAGAG	558
QY	635	CGAGAGACGCGCCATCGAGCGCGCGCGCATCTCTCTCGCGGCAACAGCTTCAGCACAGCAA	694
DB	559	AGAGAAACAATCCATCAAGCTTTCTCCCGGAGAGGTCAGCGATCTGAGTCTGCGATCTG	618
QY	695	GCGGCGCGCACAGA-----CTGCGCGCGCTCGCTACCGCGCGCGCAACGCGG	736
DB	619	GCTGAGACAGAGCGCGCGCGCGCTCCACTACTCTCTCGCGGCAACAGCTTCAGCACAGCAA	678
QY	737	GCTGCGGAGCTGATCGTCGCGCGCGCACACAGGATTCGAGGATCTTCTCGCGCGCGCGAG	796
DB	679	GCCACGCGCGCGCAAGATTCGCGCGCGCGCGAGATCGAGGCGCTTCTTCGCGCGCGCGAG	738
QY	797	GCGGCGCGCGCGCGCGCTTTGCTTCCAAAGTACAACTTCGATCTCGTCGCGCGCGCTGCGCC	856
DB	739	GAGGCTGAGGCCACAGCGCTTCGCCGCCAAGTACAACTTCGAGCTGCTTCGCGCGCTGCGCC	798
QY	857	CTCAGCGCGCGCGCGCGTTCGAGTGGCGCGCGGTCGAGCATCTGAAGC---GAGCGTGG	914
DB	799	CTCAGCGC---CGGTGCGTTCGAGTGGACTCCGGTGGTCAGCAGCGCGCAAGCTGAAGCGAG	855

QY 1288 AAGGGCGCGCTCTAGAGGATCCAGCTTAAGTACGCGTGCATGCGACGCTCATAGCTCT 1347
Db 1574 AAGGGCGCGCTCTAGAGGATCCAGCTTAAGTACGCGTGCATGCGACGCTCATAGCTCT 1633
QY 1348 TCTATAGTGTCACTAAATTCATT 1371
Db 1634 TCTATAGTGTCACTAAATTCATT 1657

RESULT 6
DRE505990
LOCUS 2658 bp mRNA linear VRT 12-AUG-2003
DEFINITION Danio rerio mRNA for heparan sulfate 6-O-sulfotransferase (hs6st gene).
ACCESSION AJ505990
KEYWORDS heparan sulfate 6-O-sulfotransferase; hs6st gene.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
AUTHORS Bink, R.J., Habu, H., Lele, Z., Dolk, E., Joore, J., Rauch, G.J., Geisler, R., Wilson, S.W., de Hertog, J., Kimata, K. and Zivkovic, D.
TITLE Heparan Sulfate 6-O-Sulfotransferase Is Essential for Muscle Development in Zebrafish
J. Biol. Chem. 278 (33), 31118-31127 (2003)
JOURNAL 12782624
PUBMED 2 (bases 1 to 2658)
REFERENCE Bink, R.J.
AUTHORS Direct Submission
TITLE Submitted (19-AUG-2002) Bink R.J., Netherlands Inst. for Dev. Biology, Hubrecht Laboratory, Uppsalalaan 8, 3584CT Utrecht, NETHERLANDS
FEATURES
source Location/Qualifiers
1. .2658
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/chromosome="14"
1. .2658
/gene="hs6st"
314..1720
/gene="hs6st"
/function="brain and fin development, myogenesis"
/codon_start=1
/product="heparan sulfate 6-O-sulfotransferase"
/protein_id="CAD44529.1"
/db_xref="GI:29150001"
/db_xref="SPTREMBL:Q800H9"
/translation="MDGKNYSKLLIALMLIFGGIVLQYICSTSDWOLLHLASLSS RLGSAPGRLNGAGADPYSSDGLAVRFPNFTKDLGRAVDHFKGDDVIVFL H1QKGTGTFVGNRIQLERCECHAGCKTCYRPGKRWTLFSRFTSGWSCGLH ADWTLNCPVSMNSRQSRMTSPRNYYITILRDPVWVYLSRWVORGATKA SHXMCGRPLTELPCYCGDDWSCGLEEFMVCYPLANNRQTLMLADLSLVGCYN LTVMSNQRMLLESKAKNLNMAFFGLTEVQKTYLFEHTFLSFAPFTLNGT RAASVEPETQRRIELNWDVLEYADLQRFQFARQERREARQRIQERRK LRAKVKSLGVGTGKAVFKPTKEPMTPEQSPFAEEKQDAEQTLESETGEVNEWLE EDDGEIMLDYSENVEQWR"

QY 1262 AAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAGCTTACGT 1321
Db 2531 AAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAGCTTACGT 2590
QY 1322 ACGCGTGCATGCGCTCATAGCTCTTCTATAGTGTCACTAAATTCATT 1371

Query Match 7.9%; Score 108.4; DB 5; Length 2658;
Best Local Similarity 99.1%; Pred. No. 1.5e-08;
Matches 109; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAGCTTACGT 1321
Db 2531 AAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAGCTTACGT 2590
QY 1322 ACGCGTGCATGCGCTCATAGCTCTTCTATAGTGTCACTAAATTCATT 1371

Query Match 7.8%; Score 107.2; DB 6; Length 2322;
Best Local Similarity 89.8%; Pred. No. 2.3e-08;
Matches 115; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1244 AAGATCTGTTGTTCTCAAAAAAAAAAAAAAAAAAAGGCGCGCTCTA 1303
Db 1830 AAGATTTAATATAGTCAAAAAAAAAAAAAAAAAAAGGCGCGCTCTA 1889
QY 1304 GAGGATCCAGCTTACGCGTGCATGCGCTCATAGCTCTTCTATAGTGTCACTA 1363
Db 1890 GAGGATCCAGCTTACGCGTGCATGCGCTCATAGCTCTTCTATAGTGTCACTA 1949
QY 1364 AATTCATT 1371
Db 1950 AATTCATT 1957

RESULT 8
BD072967
LOCUS 824 bp DNA linear PAT 27-AUG-2002
DEFINITION 70 human secretory proteins.
ACCESSION BD072967
VERSION BD072967.1 GI:22618570
KEYWORDS JP 2001514885-A/28.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ruben, S.M., Young, P.E., Brewer, L.A., Ebner, R., Olse, H.S., Florence, K.A., Rosen, C.A., Duan, R., Moore, P.A., Shi, Y., Lafleur, D.W., Florence, C., Soppet, D.R., Endress, G.A., Feng, P. and Comatsacs, G.A.
TITLE 70 human secretory proteins
JOURNAL Patent: JP 2001514885-A 28 18-SEP-2001;
COMMENT HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2001514885-A/28
PF 18-SEP-2001
PR 18-AUG-1998 JP 2000509821
PR 19-AUG-1997 US 60/056555, 19-AUG-1997 US 60/056556 PR
19-AUG-1997 US 60/056535, 19-AUG-1997 US 60/056629 PR
19-AUG-1997 US 60/056369, 19-AUG-1997 US 60/056628 PR
19-AUG-1997 US 60/056728, 19-AUG-1997 US 60/056368 PR
19-AUG-1997 US 60/056726, 16-JUN-1998 US 60/089510 PR

Db 2591 ACGGTGCATGCGCTCATAGCTCTTCTATAGTGTCACTAAATTCATT 2640

RESULT 7
AX670965
LOCUS 2322 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 3 from Patent EP1277832.
ACCESSION AX670965
VERSION AX670965.1 GI:29329461
KEYWORDS
SOURCE
ORGANISM Heliothis virescens (tobacco budworm)
REFERENCE 1 Heliothis virescens
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.
TITLE Ebens, A.J., Johnston, S. and Breach, J.C.
JOURNAL Polynucleotides encoding insect ethanalamine kinase and uses therefor
Patent: EP 1277832-A 3 22-JAN-2003;
Genoptera, LLC (US)
FEATURES
source Location/Qualifiers
1. .2322
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"

ORIGIN
Query Match 7.8%; Score 107.2; DB 6; Length 2322;
Best Local Similarity 89.8%; Pred. No. 2.3e-08;
Matches 115; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1244 AAGATCTGTTGTTCTCAAAAAAAAAAAAAAAAAAAGGCGCGCTCTA 1303
Db 1830 AAGATTTAATATAGTCAAAAAAAAAAAAAAAAAAAGGCGCGCTCTA 1889
QY 1304 GAGGATCCAGCTTACGCGTGCATGCGCTCATAGCTCTTCTATAGTGTCACTA 1363
Db 1890 GAGGATCCAGCTTACGCGTGCATGCGCTCATAGCTCTTCTATAGTGTCACTA 1949
QY 1364 AATTCATT 1371
Db 1950 AATTCATT 1957

RESULT 8
BD072967
LOCUS 824 bp DNA linear PAT 27-AUG-2002
DEFINITION 70 human secretory proteins.
ACCESSION BD072967
VERSION BD072967.1 GI:22618570
KEYWORDS JP 2001514885-A/28.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ruben, S.M., Young, P.E., Brewer, L.A., Ebner, R., Olse, H.S., Florence, K.A., Rosen, C.A., Duan, R., Moore, P.A., Shi, Y., Lafleur, D.W., Florence, C., Soppet, D.R., Endress, G.A., Feng, P. and Comatsacs, G.A.
TITLE 70 human secretory proteins
JOURNAL Patent: JP 2001514885-A 28 18-SEP-2001;
COMMENT HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2001514885-A/28
PF 18-SEP-2001
PR 18-AUG-1998 JP 2000509821
PR 19-AUG-1997 US 60/056555, 19-AUG-1997 US 60/056556 PR
19-AUG-1997 US 60/056535, 19-AUG-1997 US 60/056629 PR
19-AUG-1997 US 60/056369, 19-AUG-1997 US 60/056628 PR
19-AUG-1997 US 60/056728, 19-AUG-1997 US 60/056368 PR
19-AUG-1997 US 60/056726, 16-JUN-1998 US 60/089510 PR

```

15-JUL-1998 US 60/092956
PI STEVEN M RUBEN, PAUL E YOUNG, LAURIE A BREWER, REINHARD EBNER, PI
HENRIK S OLSEN,
PI KIMBERLY A FLORENCE, CRAIG A ROSEN, ROSANNE DUAN, PAUL A MOORE,
PI YANGGU SHI,
PI DAVID W LAFLEUR, CHARLES FLORENCE, DANIEL R SOPPET, GREGORY A PI
ENDRESS.
PI PING FENG, GEORGE A COMATISACS
PC C12N15/09, A61K38/00, A61P43/00, C07K16/00, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12Q1/68, G01N33/53, G01N33/68, C12N15/00, PC
A61K37/02,
PC C12N5/00
CC n equals a,t,g,or c
CC n equals a,t,g,or c
CC n equals a,t,g,or c
FH Key Location/Qualifiers
FT source 1..824
/organism='Homo sapiens (human)'
/organism=Homo sapiens
/mol_type='genomic DNA'
/db_xref='taxon:9606'

FEATURES
source
LOCATION/Qualifiers
1..824
/organism=Homo sapiens
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 7.8%; Score 106.4; DB 6; Length 824;
Best Local Similarity 97.3%; Pred. No. 3.2e-08;
Matches 107; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCCAAGCTTACGT 1321
DB 593 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAAGATCCCAAGCTTACGT 642

QY 1322 ACGCGTCATCGCGCTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
DB 643 ACGCGTCATCGCGCTCATAGCTCTTCTATAGTGTACCTAAATTCATT 692

RESULT 9
AX670963 2402 bp DNA linear PAT 27-MAR-2003
LOCUS Sequence 1 from Patent EP1277832.
DEFINITION AX670963
ACCESSION AX670963
VERSION AX670963.1 GI:29329460
KEYWORDS Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
SOURCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE
AUTHORS Ebens, A.J., Johnston, S. and Breach, J.C.
TITLE Polynucleotides encoding insect ethanolamine kinase and uses
therefor
JOURNAL Patent: EP 1277832-A 1 22-JAN-2003;
Genoptera, LLC (US)
FEATURES
source
LOCATION/Qualifiers
1..2402
/organism=Heliothis virescens
/mol_type='unassigned DNA'
/db_xref='taxon:7102'

ORIGIN
Query Match 7.7%; Score 105.6; DB 6; Length 2402;
Best Local Similarity 89.1%; Pred. No. 4.4e-08;
Matches 114; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1244 AAGATCTGTTTGTCTCAAAAAAAAAAAAAAAAAAGGGCGCGCTCTA 1303
DB 1939 AAGATTAAATAGTCAAAAAAAAAAAAAAAAAAGGCCATGCGCGCTCTA 1998
QY 1304 GAGGATCCAAAGCTTACGTAACGGTCATCGCATAGCTCTTCTATAGTGTACCTA 1363

```

```

DB 1999 GAGGATCCAAAGCTTACGTAACGGTCATCGCATAGCTCTTCTATAGTGTACCTA 2058
QY 1364 AATTCAATT 1371
DB 2059 AATTCAAT 2066

RESULT 10
BD276004 2454 bp DNA linear PAT 17-JUL-2003
LOCUS 62 Human secreted proteins.
DEFINITION BD276004
ACCESSION BD276004
VERSION BD276004.1 GI:33085772
KEYWORDS JP 2002543771-A/82.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2454)
AUTHORS Birse, C.E., Mouret, P.A., Florence, K.A., Ruben, S.M.,
Komatsoulis, G.A., Ni, J., Ebner, R., W.D., Lafleur, Olsen, H.S.,
Shi, Y., Soppet, D.R., Rosen, C.A. and Young, P.E.
TITLE 62 Human secreted proteins
JOURNAL Patent: JP 2002543771-A 82 24-DEC-2002;
Human Genome Sciences Inc
COMMENT OS Homo sapiens
PN JP 2002543771-A/82
PD 24-DEC-2002
PF 06-APR-2000 JP 2000611564
PR 09-APR-1999 US 60/128693, 26-APR-1999 US 60/130991 PI
Charles e birse, paul a mouret, kimberly a florence, steven m pi
ruben,
PI george a komatsoulis, jian ni, reinhard ebner, david w pi
lafleur, henrik s olsen,
PI yanggu shi, daniel r soppet, craig a rosen, paul e young cc
FH Key Location/Qualifiers
FT SITE (2317)
/notes='n equals a,t,g, or c'

FEATURES
source
LOCATION/Qualifiers
1..2454
/organism=Homo sapiens
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 7.7%; Score 105.2; DB 6; Length 2454;
Best Local Similarity 97.3%; Pred. No. 5.1e-08;
Matches 107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCCAAGCTTACGT 1321
DB 2034 AAAAAAAAAACAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCCAAGCTTACGT 2153
QY 1322 ACGCGTCATCGCGCTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
DB 2154 ACGCGTCATCGCGCTCATAGCTCTTCTATAGTGTACCTAAATTCATT 2203

RESULT 11
BD275352 1270 bp DNA linear PAT 17-JUL-2003
LOCUS 49 Human Secreted Proteins.
DEFINITION BD275352
ACCESSION BD275352
VERSION BD275352.1 GI:33085120
KEYWORDS JP 2002539787-A/28.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1270)
AUTHORS Komatsoulis, G., Rosen, C.A. and Ruben, S.M.
TITLE 49 Human Secreted Proteins

```


ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gu, W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 4 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1..2486
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<2..1903
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13077.1"
/db_xref="GI:17382884"
/db_xref="REMBL:CAD13077"
translation="NTHYRSWYACRVRSIGPGSTHASVERSGLSLPAHPASIALAAL
AASNTASGKLELDTPFSLQSLQALDLSNNAIRSIHPEAFSTLHSLVKLDLTDNOLTT
LPLAGLGLMHLKLNALQAQFSKDSFKLRIILEVYAYQCCPYGMCASFEFFKASGQ
WEADLHDDDESSRPGLLARQENHYDQDLQEDMEDSKPHPSVQCSPPTGPPF
KPCYLFESGIRLAWAIVLILVLCNGLVLLTVFAGGPAPLPVVKVVGAIAGANTL
TGICSLIASVDALTFGQFSEYGRWETGLGRATGFLIAGVSEASVILLTLIAVQCS
VSVSCVRAVGKSPSLGSGVAGVIGLALAAALPLASVGEYASPLCLFYAPPEQG
PAALGFTVALVMNSFCFLVAVAGVILKLYCDLPRGDFRAVWDCAMVRHVAWLI PADGL
LYCPVAFSLFASMLGLFPVTPENAVKSVLLVILPLACLNPLLYLLFNHPRDDLRRLR
FRAGSGPLAAAGAELEKSCDSQALVAFSDVLDLLEASGRPPGLETYGPFPSVT
LISQQPGAPRIEGSHCVPEGNHFNQPSNDGELLRLRAEGSTPAGGGLSGGGGFGFQ
SGLAPASHV"

Search completed: October 1, 2004, 23:19:49
Job time : 19996 secs

ORIGIN

Query Match 7.1%; Score 97; DB 6; Length 2486;
Best Local Similarity 95.2%; Pred. No. 1.3e-06;
Matches 100; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1267 AAAAAAAAAAAAAAAAAAAGGCGCGCCCTCTAGAGGATCCAAGCTTACGTACGCG 1326
Db 2381 AAGAGATTAAAAAAAAAAAAAAAAAGGCGCGCCCTCTAGAGGATCCAAGCTTACGTACGCG 2440
QY 1327 TGCATGCGACGCATAGCTCTTTCTATAGTGTCACTTAATTCATT 1371
Db 2441 TGCATGCGACGCATAGCTCTTTCTATAGTGTCACTTAATTCATT 2485

RESULT 15

AR066494
LOCUS 1581 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5850022.
ACCESSION AR066494
VERSION AR066494.1 GI:5996710
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1581)
AUTHORS Dehesh,K., Voelker,T. and Hawkins,D.
TITLE Production of myristate in plant cells
JOURNAL Patent: US 5850022-A 1 15-DEC-1998;
FEATURES
source 1..1581
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 7.0%; Score 95.6; DB 6; Length 1581;
Best Local Similarity 96.1%; Pred. No. 2.2e-06;
Matches 98; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1270 AAAAAAAAAAAAAAAAAAAGCGCGCCCTCTAGAGGATCCAAGCTTACGTACGCGTCG 1329
Db 1474 AAGTTAAAAAAAAAAAAAAAAAGCGCGCCCTCTAGAGGATCCAAGCTTACGTACGCGTCG 1533

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 17:12:11 ; Search time 453 Seconds
(without alignments)
1680.778 Million cell updates/sec

Title: US-09-993-808B-1

Perfect score: 1372

Sequence: 1 cccacgcgtccgacgcaag.....agtgtcacctaaattcattc 1372

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCRU COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	95.6	7.0	1581	1	US-08-383-756-1
2	95.6	7.0	1581	2	US-08-460-898-1
3	95.6	6.9	1255	3	US-09-118-442-31
4	95.6	6.9	1255	3	US-09-677-064-31
5	93.2	6.8	1384	4	US-09-904-615-46
6	91.6	6.7	1665	4	US-09-591-095-23
7	91.6	6.6	944	4	US-09-227-357-122
8	88.8	6.5	2581	2	US-09-013-634-1
9	84.6	6.1	696	4	US-09-227-357-69
10	82.4	6.0	831	4	US-09-904-615-25
11	82.0	6.0	1330	3	US-09-118-442-29
12	82.0	6.0	1330	3	US-09-677-064-29
13	79.4	5.8	826	4	US-09-620-312D-584
14	79.4	5.8	907	4	US-09-620-312D-585
15	79.4	5.8	1200	4	US-09-620-312D-564
16	79.4	5.8	2544	4	US-09-483-371-1
17	79.4	5.8	2544	4	US-09-957-156-1
18	79.4	5.8	955	4	US-09-620-312D-228
19	78.8	5.7	2638	3	US-09-042-785A-22
20	78.4	5.7	501	4	US-09-439-313-451
21	78.4	5.7	501	4	US-09-352-616A-451
22	78.4	5.7	501	4	US-09-636-215-451
23	78.4	5.7	501	4	US-09-685-166A-451
24	78.4	5.7	706	4	US-09-439-313-449
25	78.4	5.7	706	4	US-09-352-616A-449
26	78.4	5.7	706	4	US-09-636-215-449
27	78.4	5.7	706	4	US-09-685-166A-449

28	78.4	5.7	1663	4	US-09-620-312D-6
29	78.4	5.7	1790	4	US-09-118-637A-3
30	77.8	5.7	1243	4	US-09-620-312D-1092
31	77.8	5.7	1006	3	US-08-911-423-3
32	77.8	5.7	3327	3	US-08-689-421-26
33	77.8	5.7	3327	3	US-09-389-528-26
34	77.8	5.7	3327	3	US-09-181-827A-26
35	77.4	5.6	2311	4	US-09-489-847-123
36	77.4	5.6	1689	4	US-09-053-374A-4
37	76.6	5.6	1067	4	US-09-620-312D-543
38	76.4	5.6	1067	4	US-09-620-312D-853
39	76.4	5.5	2085	4	US-09-620-312D-1002
40	76.4	5.5	2664	4	US-09-149-476-255
41	76.4	5.5	3396	4	US-09-668-680-6
42	76.4	5.5	3423	4	US-09-668-680-7
43	75.5	5.5	2202	4	US-09-396-149-3
44	74.8	5.5	322	4	US-08-956-171E-1520
45	74.2	5.4	2727	4	US-09-620-312D-218

ALIGNMENTS

RESULT 1
US-08-383-756-1
; Sequence 1, Application US/08383756
; Patent No. 5654495
; GENERAL INFORMATION:
; APPLICANT: Dehesh, Katayoon
; APPLICANT: Voelker, Toni Alois
; APPLICANT: Hawkins, Deborah
; APPLICANT: Davies, Huw Maelor
; TITLE OF INVENTION: Production of Myristate in Plant Cells
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,756
; FILING DATE: 02-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10814
; FILING DATE: 23-OCT-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/968,971
; FILING DATE: 30-OCT-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:

Sequence 6, Appli
Sequence 3, Appli
Sequence 1092, Ap
Sequence 3, Appli
Sequence 26, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 123, App
Sequence 4, Appli
Sequence 543, App
Sequence 853, App
Sequence 1002, Ap
Sequence 255, App
Sequence 6, Appli
Sequence 7, Appli
Sequence 3, Appli
Sequence 1520, Ap
Sequence 218, App

SEQUENCE CHARACTERISTICS:
LENGTH: 1581 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-383-756-1

Query Match 7.0%; Score 95.6; DB 1; Length 1581;
Best Local Similarity 96.1%; Pred. No. 1.7e-11;
Matches 98; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1270 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 1474 AAGTTAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1533

Qy 1330 ATGGAGCTCATAGCTCTTCTATAGTGTCACTAAATTCAT 1371
Db 1534 ATGGAGCTCATAGCTCTTCTATAGTGTCACTAAATTCAT 1575

RESULT 2
US-08-460-898-1
Sequence 1, Application US/08460898
Patent No. 5850022
GENERAL INFORMATION:
APPLICANT: Denesh, Katayoon
APPLICANT: Voelker, Toni Alois
APPLICANT: Hawkins, Deborah
APPLICANT: Davies, Huw Maelor
TITLE OF INVENTION: Production of Myristate in Plant Cells
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1(a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,898
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/383,756
FILING DATE: 02-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,695
FILING DATE: 16-JUN-94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10814
FILING DATE: 29-OCT-93
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 30-OCT-92
FILING DATE: 30-OCT-92
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 111-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313

TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1581 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-460-898-1

Query Match 7.0%; Score 95.6; DB 2; Length 1581;
Best Local Similarity 96.1%; Pred. No. 1.7e-11;
Matches 98; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1270 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 1474 AAGTTAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1533

Qy 1330 ATGGAGCTCATAGCTCTTCTATAGTGTCACTAAATTCAT 1371
Db 1534 ATGGAGCTCATAGCTCTTCTATAGTGTCACTAAATTCAT 1575

RESULT 3
US-09-118-442-31
Sequence 31, Application US/09118442B
Patent No. 6197561
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
APPLICANT: Wang, Xun
APPLICANT: Bowen, Benjamin A.
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
TITLE OF INVENTION: Plants and Uses Thereof
FILE REFERENCE: 0706
CURRENT APPLICATION NUMBER: US/09/118,442B
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/055,446
EARLIER FILING DATE: 1997-08-11
EARLIER APPLICATION NUMBER: 60/055,526
EARLIER FILING DATE: 1997-08-08
EARLIER APPLICATION NUMBER: 60/053,944
EARLIER FILING DATE: 1997-07-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 1255
TYPE: DNA
ORGANISM: Zea mays
US-09-118-442-31

Query Match 6.9%; Score 95; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1275 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1334
Db 1161 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1220

Qy 1335 ACGTCATAGCTCTTCTATAGTGTCACTAAATTCAT 1369
Db 1221 ACGTCATAGCTCTTCTATAGTGTCACTAAATTCAT 1255

RESULT 4
US-09-677-064-31
Sequence 31, Application US/09677064
Patent No. 6291224
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.

; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE OF INVENTION: Plants and Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-31

Query Match 6.9%; Score 95; DB 3; Length 1255;
Best Local Similarity 100.8%; Pred. No. 2.1e-11;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1275 AAAAAAAAAAAAAAGGCGCGCCTCTAGAGGATCCAAAGCTTACGTACGGTGCGTGCATGCG 1334
Db 1161 AAAAAAAAAAAAAAGGCGCGCCTCTAGAGGATCCAAAGCTTACGTACGGTGCGTGCATGCG 1220
QY 1335 ACCTCATAGCTTCTTATAGTGTCACCTAAATCA 1369
Db 1221 ACCTCATAGCTTCTTATAGTGTCACCTAAATCA 1255

RESULT 5

US-09-904-615-46
; Sequence 46, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 1984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-46

Query Match 6.8%; Score 93.2; DB 4; Length 1984;
Best Local Similarity 85.2%; Pred. No. 5.9e-11;
Matches 104; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1240 AAGAAAGATCTGGTTGTCTCAAAAAAATAACACACAAAAAAGGCGCGCGC 1299
Db 1863 AAAAAAACAATTTACCTTAAGAAATAACACACAAAAAAGGCGCGCGC 1922
QY 1300 TCTAGAGGATCCAAAGCTTACGTACGGTGCGTACGCTTCTTATAGTGTC 1359
Db 1923 TCTAGAGGATCCAAAGCTTACGTACGGTGCGTACGCTTCTTATAGTGTC 1982

QY 1360 CC 1361
Db 1983 CC 1984

RESULT 6

US-09-591-095-23
; Sequence 23, Application US/09591095
; Patent No. 6489461
; GENERAL INFORMATION:
; APPLICANT: Proman, B. & Dehesh, K.
; TITLE OF INVENTION: Nucleic Acid Sequences encoding Proteins
; TITLE OF INVENTION: Involved in Fatty Acid Beta-Oxidation and Methods of Use.
; FILE REFERENCE: 17137/01/US
; CURRENT APPLICATION NUMBER: US/09/591,095
; CURRENT FILING DATE: 2000-06-08
; EARLIER APPLICATION NUMBER: US 60/138,162
; EARLIER FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Brassica sp.
US-09-591-095-23

Query Match 6.7%; Score 91.6; DB 4; Length 1665;
Best Local Similarity 95.5%; Pred. No. 1.2e-10;
Matches 105; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1262 AAAAAAAAAAAAAAGGCGCGCCTCTAGAGGATCCAAAGCTTACGT 1321
Db 1545 AAAAAAAAAAACCAAAAAAAGGCGCGCCTCTAGAGGATCCAAAGCTTACGT 1604
QY 1322 ACGGTGCATGCGACGTCATAGCTTCTTATAGTGTCACCTAAATTCATT 1371
Db 1605 ACGCG-GCATGCGACGTCATAGCTTCTTATAGTGTCACCTAAATTCATT 1653

RESULT 7

US-09-227-357-122
; Sequence 122, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 122
LENGTH: 944
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (932)
FEATURE:
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (942)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (944)
OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-122

Query Match 6.6%; Score 91; DB 4; Length 944;
Best Local Similarity 68.5%; Pred. No. 1.4e-10;
Matches 139; Conservative 1; Mismatches 61; Indels 2; Gaps 1;
QY 1149 AATCCGCTTACTGTGTAATTTCTCTCATGATCTCTTAGATGAGATTATCGTCTTAAT 1208
DB 724 AAGCAATTAGTGAAGCACTTCTATCCAAATGACCTTTTGTCCCTTTTAAACCAAT 783
QY 1209 TTATTAAGTACGAGCTGTGGCTTCCAAAGAAAGATCTGGTTTGTCTCAAAAAA 1268

Db 784 TTACT--GTTACTGAAACTTTTGTACATAAAGCAATACGAGATTAAGAAAAA 841
QY 1269 AAAAAAAGGGCGGCTCTAGAGATCCAAAGCTTAGTAGCGGTG 1328
|||||
Db 842 AAAAAAAGGGCGGCTCTAGAGATCCAAAGCTTAGTAGCGGTG 901
QY 1329 CATGCGAGCTCATAGCTTCTTA 1351
|||||
Db 902 CATGCGAGCTCATAGCTTCTTA 924
RESULT 8
US-09-013-634-1
Sequence 1, Application US/09013634
Patent No. 5945307
GENERAL INFORMATION:
APPLICANT: M. Alexandra Glucksmann and Keith Robinson
TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,634
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Attorney, Jean M. Silveri
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 184..1194
US-09-013-634-1

Query Match 6.5%; Score 88.8; DB 2; Length 2581;
Best Local Similarity 88.9%; Pred. No. 5.5e-10;
Matches 96; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1241 AAGAAAGATCTGTTTGTCTCAAAAAAAGGGCGGCGGT 1300
Db 2471 AATAAAACCAAGGAGAGCAAAAAAAGGGCGGCGGT 2530
QY 1301 CTAGAGGATCCAGCTTAGTAGCGGTGATCGAGCTCATAGCTCTT 1348
Db 2531 CTAGAGGATCCAGCTTAGTAGCGGTGATCGAGCTCATAGCTCTT 2578

RESULT 9
US-09-227-357-69
Sequence 69, Application US/09227357

```

; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/055,661
; EARLIER FILING DATE: 1997-09-12

```

```

; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 69
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (605)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (648)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (655)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-69

Query Match 6.1%; Score 84; DB 4; Length 696;
Best Local Similarity 92.2%; Pred. No. 4e-09;
Matches 95; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1270 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGC 1329
Db 567 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAKAGGNTCCAGCTTACGTACGGTGC 626

QY 1330 ATGCGACGTCATAGCTCTTCTAT-AGTGTACCACTAAATTCATT 1371
Db 627 ATGCGACGTCATAGCTCTTCTNTAAGGNCACCTAAATTCAT 669

```

```

RESULT 10
US-09-904-615-25
; Sequence 25, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 25
; LENGTH: 831
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-25

```


QY 1351 ATAGTGTACCTAAATTCATT 1371
Db 803 ATAGTGTACCTAAATTCATT 823

RESULT 14

US-09-620-312D-585
; Sequence 585, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: Pt_FL_genes Version 1.0
; SEQ ID NO 585
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)...(695)
US-09-620-312D-585

Query Match 5.8%; Score 79.4; DB 4; Length 907;
Best Local Similarity 98.8%; Pred. No. 4.2e-08;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1291 GCGGGCGCTCTAGAGATCCAAAGCTTAGCTAGCGGTGATGGAGTCATAGCTCTTCT 1350
Db 824 GCGGGCGCTCTAGAGATCCAAAGCTTAGCTAGCGGTGATGGAGTCATAGCTCTTCT 883
QY 1351 ATAGTGTACCTAAATTCATT 1371
Db 884 ATAGTGTACCTAAATTCATT 904

RESULT 15

US-09-620-312D-564/c
; Sequence 564, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: Pt_FL_genes Version 1.0
; SEQ ID NO 564
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (559)...(957)
US-09-620-312D-564

Query Match 5.8%; Score 79.4; DB 4; Length 1200;
Best Local Similarity 98.8%; Pred. No. 4.5e-08;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1291 GCGGGCGCTCTAGAGATCCAAAGCTTAGCTAGCGGTGATGGAGTCATAGCTCTTCT 1350
Db 86 GCGGGCGCTCTAGAGATCCAAAGCTTAGCTAGCGGTGATGGAGTCATAGCTCTTCT 27
QY 1351 ATAGTGTACCTAAATTCATT 1371
Db 26 ATAGTGTACCTAAATTCATT 6

Search completed: October 2, 2004, 05:36:30
Job time : 458 secs

This Page Blank (uspto)

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.
Location/Qualifiers
1. .707

FEATURES

source

ORIGIN

Query Match.	42.4%;	Score 581.8;	DB 11;	Length 707;
Best Local Similarity	86.7%;	Pred. No. 1.1e-58;		
Matches 615;	Conservative 1;	Mismatches 80;	Indels 13;	Gaps 1;
QY	525	GCGGAGCCACACTGCGGTCCTACAGACGCTGCAGAGGCTGGCGGGATCATACGCTCTCGG	584	
Db	8	GCGGAGCCACACTGCGGTCCTACAGACGCTGCAGAGGCTGGCGGGATCATACGCTCTCGG	67	
QY	585	TGGATGTCCTGGCGGCGAGCAACTCCGGGAGCGGCCAGACCGCGAGAGCGGAGAGACGA	644	
Db	68	TGGTGGATGTCCTGGCGGAGCAACTCCGGGAGCGGCCAGACCGCGAGAGAGAGACGA	127	
QY	645	CGCCATCGAGCGGGCGCACGGCGAGCTCAGCGATCTGGAGTTCGAGTCTGGCGGGGACAC	704	
Db	128	CGCCATCGAGCGGGCGCACGGCGAGCTCAGCGATCTGGAGTTCGAGTCTGGCGGGGACAC	187	
QY	705	AGACTGGCCCGTCCGTACCCGCGGCAAGCGCGCTGCGAGCTCATCGTCCGCCGACGAC	764	
Db	188	AGACTGGCCCGTCCGTACCCGCGGCAAGCGCGCTGCGAGCTCATCGTCCGCCGACGAC	247	
QY	765	ACGAGATCCAGGAGTTCCTTCGCCCGCGCGAGGCGGCCAGGCGCTTTCGCTTCCA	824	
Db	248	ACGAGATCCAGGAGTTCNN	307	
QY	825	AGTACAACTTCGACTTCGTCCGCGCGTGCCTTCGACGCGCGCGCGGTCGAGTGGG	884	
Db	308	AGTACAACTTCGACTTCGTCCGCGCGTGCCTTCGACGCGCGCGCGGTCGAGTGGG	367	
QY	885	CGCCGCTGTCGAGATCTGAAGCAGAGGTGGGTCCGCTGCAAGTGAAGCTAGAAAGAGA	944	
Db	368	CGCCGCTGGTCGAGATCTGAAGCAGAGGTGGGTCCGCTGCAAGTGAAGCTAGAAAGAGA	427	
QY	945	AAAGATGCCCCCCCCCCCCCCCCCAACAACATAACGGAAGAGAAAAACCAACAAT	1004	
Db	428	AAAGATG-----NNNNNAAACAACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	474	
QY	1005	TAAGCAGCTTTATATAGCCTAAGCTAACCAACCAACCAATTCATCTCGTCCAAATGCATGCCT	1064	
Db	475	NAAGCAGCTTTATATAGCCTAAGCTAACCAACCAACCAATTCATCTCGTCCAAATGCATGCCT	534	
QY	1065	TGCTTTTCTGAGAGCTAGCAGGAGCGTAGTTATATTTAGTACTACTTTACTTTATTCAG	1124	
Db	535	TGCTTTTCTGAGAGCTAGCAGGAGCGTAGTTATATTTAGTACTACTTTACTTTATTCAG	594	
QY	1125	AGGTTATCTTGACCCGATAGATCAATCCGGTACTGTGTAATTTCTCATGATCTCT	1184	
Db	595	AGGTTATCTTGACCCGATAGATCAATCCGGTACTGTGTAATTTCTCATGATCTCT	654	
QY	1185	TAGATGGAGTTTAAATCGTCTTAAATTTATTTACTGTACAGACGCTTGSTTG	1233	
Db	655	TAGATGGAGTTTAAATCGTCTTAAATTTATTTACTGTACAGACGCTTGCTGG	703	

RESULT 2

CD444277

STUDY LOCUS

DEFINITION

ACCESSION
DEFINITION

ACCESSION
VEDSTON

NOTES
REVIEWS

KEYWORDS
CONTRACT

SOURCE
ORGANIZATION

Zea mays

Eukaryotes

Spermatop

clade; Pa

1, (bases

Lai, J., I

Messing, C.

Sequencing

Unpubl ish

Contact:

Dr. Joach

Waksman :

190 Frel:

Tel: 732-

Fax: 732.

Email: j...

Seq prime

44

19

ORIGIN

Query Match	36.1%	Score 495.6	DB 14	Length 526	
Best Local Similarity	99.2%	Pred. No. 1.3e-48			
Matches	498	Conservative 0	Mismatches 4	Indels 0	Gaps 0
Qy	13	GAGCAAGCGGTCAGGACAGCAGCGCGCGCAGGCGCTGTGGCGCTGTGGAGAGGAAAA	72		
Db	25	GAGCAAGCGGTCAGGACAGCAGCGCGCGCAGGCAATTGTGGCGCTGTGGAGAGGAAAA	84		
Qy	73	AGAAAAAGAGAAACCGGCCCAAGACAAGCAAGAGAGCCAGGGCCGCGCGCTTGCCTGCA	132		
Db	85	GGAGAAAGAGAACCGGCCCAAGACAAGCAAGAGAGCCAGGGCCGCGCGCTTGCCTGCA	144		
Qy	133	GATGGGAAATACATGTCGCAAGTCCAGGGGCGCCAGAGCGCGAGAGTCCGCCGCTCGA	192		
Db	145	GATGGGAAATACATGTCGCAAGTCCAGGGGCGCCAGAGTCCAGGGCGCGAGGTCCGCCGCTCGA	204		
Qy	193	GGTTACGACGGTGGTCGGCGTCCGGAAGAGTCCAGGTCCCGGCGCGGACACGGCGGGT	252		
Db	205	GGTTACGACGGTGGTCGGCGTCCGGAAGAGTCCAGGTCCCGGCGCGGACACGGCGGGT	264		
Qy	253	CGCGAAGTCCGCCCGAGGAGGAAGAGGCGCGCGCGGGGAGGCTGTCTCGCGCGTGAG	312		
Db	265	CGCGAAGTCCGCCCGAGGAGGAAGAGGCGCGCGCGGGGAGGCTGTCTCGCGCGTGAG	324		
Qy	313	CGCTGGTGGGACGCGCGAAGCTGCTATCATCCACTCGGTAGCCGCGATGTGTATGGC	372		
Db	325	CGCTGGTGGGACGCGCGAAGCTGCTATCATCCACTCGGTAGCCGCGATGTGTATGGC	384		
Qy	373	ACCGCCTCAGCCGACGCGTCCGTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGG	432		
Db	385	ACCGCCTCAGCCGACGCGTCCGTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGG	444		
Qy	433	CGCTGCAGGACAGCAGGCGCGCGGCTCCGCGCGCGGCTCTCGCGTTCCTCAGCACGCG	492		
Db	445	CGCTGCAGGACAGCAGGCGCGCGGCTCCGCTCCCGGGCTCTCGCGTTCCTCAGCACGCG	504		
Qy	493	GTCTCGGTTGAACTTGGGCTTG	514		

```

Db      505 GTGCTGTGGAAGTGGGCTTG 526
||||| ||||| ||||| ||||| |||||
RESULT 3
AY108168
LOCUS   Zea mays PC0119368 mRNA sequence.
DEFINITION
ACCESSION AY108168
VERSION   AY108168.1 GI:21211246
KEYWORDS HTC.
SOURCE   Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1197)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 1197)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
1..1197
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:636945"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
ORIGIN
Query Match 33.2%; Score 455.2; DB 11; Length 1197;
Best Local Similarity 71.5%; Pred. NO. 4.7e-44;
Matches 851; Conservative 0; Mismatches 233; Indels 106; Gaps 15;
QY 22 GGCTGAGGACGAGCGCGCGAGCGGTGTGGCGCTGTGGAGAGAGAAAGAGAAAG 81
Db 37 GCTGTGAGGACGCGCGCGAGCGGTGGGAGTGGCGAGTGGAGTGGAGTGAAGAA 96
QY 82 GGAAACCGGCGAAGCAAGCAAGCAAGCGAGGCGAGG-----GCCGCGCGTGGCTCAGAT 135
Db 97 GGAAACCGGCGAAGCAAGCAAGCAAGCAAGCGAGTGGCTGGCGCGCGTTCGGTAAGAT 156
QY 136 GGGGAAGTACATCGCAAGTGCAGGGGCGCGCGAGGCGGAGGTCCGCGCGCTCAGGT 195
Db 157 GGGGAAGTACATCGCAAGCGAGGCGGCGCGCGAGGCGGAGGTGGCGCGAGTCCAGGT 216
QY 196 TAGCGAGGTGCTGGCGGTCCGAGACGAGTCCAGTCCGCGCGCGCGAC---CGCGGTGT 252
Db 217 CTGCGAGGTGCTGGCGGTCCGAGACGAGTCCAGTCCGCGCGCGCGACCGCGCGCGGTGT 276
QY 253 CGCGAAGGTGCGCGCGCGAGGAGGA-----GGCGCGC 285
Db 277 CGCGAAGGTGCTGCGCGCGAGGAGGAAGGCGGTGCTGCGCGCGCGAAGCGTACGAC 336

```

```

QY 286 GCGCGGGAGCCCTGCTGCGCGCGCTGAGCGCTGCTGGGAGCGCGAA---GCTGTACAT 342
Db 337 GTGCGGGAGCCCTGCTGCGCGCTGGGCGCTGGTGGTGGGAGCGCGAAGTGTGCTACAT 396
QY 343 CCACCTGCGCTAGCCGATGCTGTTATGTCACCGCTCAGCGCGCAGCGCGTGGTGAAGT 402
Db 397 CCACCTGCGGAGCGCATGCTGTTATGTCAGCACCTCAGCAGCAACCGTGG-----C 450
QY 403 GGTTCGACCCCGGTGGAGGTGCTGATGCGCTGTCAGGACAGCAGGCGCGCGTTCGC 462
Db 451 GGTCTGACCGCGGTGGAGGTGCTGCTGCGGCACAGCAGGCGCGGTGGTGGCGTGC 510
QY 463 GCGCGGCTCTGCGGTGCTCCAGCACGCGCTGCTGCTGTAACCTGGGCTTGGGGGTCA 522
Db 511 GGTGCGCTCTGCGGTGCTCCAGCACGCGCTGCTGCTGGAGCT----- 555
QY 523 GCGCGGAGCCACACCTGCGCTTCTTACGACGCTGTCAGAGGCTGGCGGGATCACGTCT 582
Db 556 ---CGGGGGCCAGCTGCGCTTCCGACGCTGGGCTGGAGTTGACGGGGATCACGT 612
QY 583 GGTGATGCTGCGCGCGAGCAACTCCGGGAGCGGCCAGACCGCGAGGCGAGAGAC 642
Db 613 CCGGATGCTGTCACCGCGAGCAACTCCGGGAGCGTCCCGGACCGCGAGAGAGAGAC 672
QY 643 GAGGCCA---TCGAGCGCGGCGCA---CGGCGAGCTCAGGATCTGAGTTCGATCTGGC 696
Db 673 GAGCGCATGTCGAGCGCGCGCGACCGCGCGAGCTCAGCGATCTGGAGTTCGATCTGGT 732
QY 697 GG---GGCAAGACTGGC-----CCGTGCTTACCGCGGCAACCGCGCTCGCGAGCT 747
Db 733 GGGCGGCGAGAGACTGGCTGCTGCTGCTGCGCGGCAACAACATCGGTTTCGGAGCT 792
QY 748 GATGTCGCGCGCAGACACGAGATCCAGAGTTCTTTCGCGCGCGCGAGCGCGCCAGGC 807
Db 793 GATGTCGCGCGCAGCAGAGATCCAGGAATTTCTTCGCGCGCGCGAGCGCGCCATGC 852
QY 808 CAGCGCTTGTGCTTCAAGTACAACTTCGACTTCGTCGCGCGCGTCCCTTCAGCGCGG 867
Db 853 CAAACGCTTGTGCTTCAAGTACAACTTCGACTTCGTCGCGCGCGTCCCTTCAGCGC--- 909
QY 868 CGCGCGCTTTCGAGTGGCGCGCTGCTCAGCATCTGAAGCGAGCGTGGC-----TCGG 921
Db 910 CGGCGGTTTCGAGTGCAGCGCGAGGTGCTGAGCATCTGAAGCGAGCGTGGTGCAGGTGA 969
QY 922 TGCAGGTGAGCTGAGAGAGAAAG-ATGCCCCCCCCCCCCCCCCCCCCCAACAAATAA 980
Db 970 AGCTACTACTAGTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1029
QY 981 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1040
Db 1030 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084
QY 1041 TTTCATCTGCTCCAAATGCTGCTTTCCTGAGCTAGCAGGAGCTAGTATTATTA 1100
Db 1085 AACATC-----TGCTTGTCTTTTCTCCGAGTAGACTTGTAGATGACAG 1131
QY 1101 TTGATCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1150
Db 1132 CGTAGTATTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1181

```

```

RESULT 4
BZ743686/c
LOCUS   BZ743686 869 bp DNA linear GSS 03-MAR-2003
DEFINITION
OGFAC93TC ZM 0.7 1.5 KB Zea mays genomic clone ZMM8Ma0228P17,
genomic survey sequence.
ACCESSION BZ743686
VERSION   BZ743686.1 GI:28723396
KEYWORDS GSS.
SOURCE   Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```


QY 429 ATGCGCTGACAGACAGAGCGCGCGCTCGCGCGCGGCTCTCGCTTGTCTCCAGCA 488
 Db 175 ATGCGCTGACAGACAGAGCGCGCTCTCGCGCGCGGCTCTCGCTTGTCTCCAGCA 116
 QY 489 CGGCGTCTGCGTGAATCTTGGGCTTGGGCGGTGACGCGGAGCCACACCTGCGCTCCT 548
 Db 115 CGGCGTCTGCGTGAATCTTGGGCTTGGGCGGTGACGCGGAGCCACACCTGCGCTCCT 56
 QY 549 ACGACGCTGCAGAGG 563
 Db 55 ACGACGCTGCAGAGG 41

RESULT 6
 CC440801/c
 LOCUS PHLN53TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBTa478J10,
 DEFINITION genomic survey sequence.
 CC440801
 VERSION CC440801.1 GI:30942238
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 838)

White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.N., Yuan, Y., San Miguel, P., Ma, J. and
 Bennetzen, J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other GSSs: PHLN53TD
 Contact: Cathy Whitelaw
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source
 1..838
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMBTa478J10"
 /clone_lib="ZM_0.6_1.0_KB"
 /notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
 Cot selected genomic DNA library"

ORIGIN
 Query Match 31.5%; Score 432.8; DB 28; Length 838;
 Best Local Similarity 95.6%; Pred. No. 2.2e-41;
 Matches 480; Conservative 0; Mismatches 12; Indels 10; Gaps 3;

QY 13 GAGCAAGCGGCTGACGAGCAGCGCGCGCGGCTTGTGACC-TGTGGAGAGGAAA 71
 Db 512 GAGCAAGCGGCTGAGAGGAGCAGCGCGCGCGGCTTGTGACC-TGTGGAGAGGAAA 453
 QY 72 AAGAGAAAGAGGAAACCGGCCACAGCAAGCAGAGGCGCGCGGCTTGTGCTC 131
 Db 452 AAGAGAAAGAGGAAACCGGCCACAGCAGAGGCGCGCGGCTTGTGCTC 393
 QY 132 AGATGGGAGTACATGCGCAAGTGCAGGGGCGCGCGGAGGTGCGCGCGCTG 191
 Db 392 AGATGGGAGTACATGCGCAAGTGCAGGGGCGCGCGGAGGTGCGCGCGCTG 333
 QY 192 AGTTACGAGGTCGTCGGCTCCGACGAGTTCAGGTCCGCGCGGCGGACCGCGGTT 251
 Db 332 AGTTACGAGGTCGTCGGCTCCGACGAGTTCAGGTCCGCGCGGCGGACCGCGGTT 273

QY 252 TCGGAGAG---GTCCGCCCGAGGAGGAGGCGCGCGCGGAGGCTTGTCTCCCGCG 308
 Db 272 TCGGAGAGGTCGTGCGCCCCGAGGAGGAGGCGCGCGCGGAGGCTTGTCTCCCGCG 213
 QY 309 TGAGCGCTGCTGGGAGCGCGGAGGCTGTACATCCACCTGCGTAGCGCATGCTGTTC 368
 Db 212 TGGGCGCTGCTGGGAGCGCGGAGGCTGTACATCCACCTGCGTAGCGCATGCTGTTC 153
 QY 369 TGGACCGCTCAGCGCGCAGCGCTCGGTGACTCGGTTCGACCCCGGAGGCTGCTG 428
 Db 152 TGGACCGCTCAGCGCGCAGCGCGG-----TCGGTTCGACCCCGGAGGCTGCTG 99
 QY 429 ATGCGCTGCGAGACAGCAGCGCGCGCTCGCGCGCGGCTCTCGGTTCGAGCA 488
 Db 98 ATGCGCTGCGAGACAGCAGCGCGCGCTCGCGCGCGGCTCTCGGTTCGAGCA 39
 QY 489 CGGCGCTGCTGCGTGAACCTTGGG 510
 Db 38 CGGCGCTGCTGCGTGAACCTTGGG 17

RESULT 7
 CA402478/c
 LOCUS EL01N0438A02.g_EndospERM_4_Zea_mays_mRNA_linear_EST_07-NOV-2002
 DEFINITION
 ACCESSION CA402478
 VERSION CA402478.1 GI:24767334
 KEYWORDS EST.
 SOURCE Zea mays

ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 505)

Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
 Messing, J.
 Sequencing of the maize endospERM ESTs
 Unpublished (2002)
 Contact: Lai, Jinsheng
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA

Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T7.
 Location/Qualifiers
 1..505
 /organism="Zea mays"
 /mol_type="mRNA"
 /culturvar="W22"
 /db_xref="taxon:4577"
 /tissue_type="EndospERM of 7-23DAP"
 /clone_lib="EndospERM 4"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

FEATURES
 source
 1..505
 /organism="Zea mays"
 /mol_type="mRNA"
 /culturvar="W22"
 /db_xref="taxon:4577"
 /tissue_type="EndospERM of 7-23DAP"
 /clone_lib="EndospERM 4"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN
 Query Match 28.6%; Score 392.8; DB 14; Length 505;
 Best Local Similarity 92.2%; Pred. No. 1.2e-36;
 Matches 437; Conservative 1; Mismatches 28; Indels 8; Gaps 2;

QY 816 TTGCTTCCAGTACAACTTCGACTTCGTCCGGGGGTGCCCTCGACGCGCGCGCGT 875
 Db 505 TTGCTTCCAGTACAACTTCGACTTCGTCCGGGGGTGCCCTCGACGCGCGCGCGT 446
 QY 876 TCGAGTGGCGCGGTGGTGCAGCTCTGAAGCAGCGGTCCGTCGAGGTGAGCT 935
 Db 445 TCGAGTGGCGCGGTGGTGCAGCTCTGAAGCAGCGGTCCGTCGAGGTGAGCT 386
 QY 936 AGAAAGAGAGAAAGATGCCCCCCCCCCCCCCCAACATAACGAGAGAGAGAAAA 995
 Db 385 AGAAAGAGAGAAAGATG-----CCCCCCCCCCCCCAACATAACGAGAGAGAGAAAC 333

QY	996	CCAAACAAATTAAAGCAGCTTTATATAGCCTAAGCTAAGCAACCAACCAATTCATCTCGTCCAA	1055
Db	332	CCAAACAAATTAAAGCAGCTTTATATAGCCTAAGCTAAGCAACCAACCAATTCATCTCGTCCAA	273
QY	1056	TGCATGCTTGCCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATATTAGTACTACTTTA	1115
Db	272	TGCATGCTTGCCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATATTAGTACTACTTTA	213
QY	1116	CTTATTCAGAGGTATCTTCACCCCGATAGATCAATCCGCTTACTGTGTAAATTTCTCTCA	1175
Db	212	CTTATTCAGAGGTATCTTCACCCCGATAGATCAATCCGCTTACTGTGTAGTTTGTCTCA	153
QY	1176	TGCATCTCTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCACTTGSTTGGC	1235
Db	152	TGCATCTCTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCACTTGC-TGGC	94
QY	1236	TTGCAAGAGAAGATCTGGTTGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAA	1289
Db	93	TTGCAAGAGAAGATCTGGTTGTCTCTTAATCCGAGTCACGTTTAGCTTAAAAAA	40

RESULT 8	CA074449	659 bp	mRNA	linear	EST 23-SEP-2003			
LOCUS	SCEZAM1081D06.g	AM1	Saccharum officinarum	cdna clone	SCEZAM1081D06			
DEFINITION	5', mRNA sequence.							
ACCESSION	CA074449							
VERSION	CA074449.1	GI:34926722						
KEYWORDS	EST.							
SOURCE	Saccharum officinarum							
ORGANISM	Saccharum officinarum							
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.							
AUTHORS	1 (bases 1 to 659)							
TITLE	Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.							
JOURNAL	The libraries that made SUCEST							
COMMENT	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)							
	Contact: Arruda, p							
	Centro de Biologia Molecular e Engenharia Genetica							
	Universidade Estadual de Campinas							
	Caixa Postal 6010, 13083-970, Campinas SP, Brazil							
	Tel: 55 19 3788 1137							
	Fax: 55 19 3788 1089							
	Email: parruda@unicamp.br							
	Clone distribution: clone distribution information can be found							
	through the Brazilian Clone Collection Center (BCCC) at							
	http://www.bccccenter.fcav.unesp.br							
	plate: 081 row: D column: 06							
	Seq primer: T7 Promoter Primer.							

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	

RESULT	9
CG324875/c	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

CG324875 885 bp DNA linear GSS 26-AUG-2003
 OGXND04TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0561H08,
 genomic survey sequence.
 CG324875
 CG324875.1 GI:34242141
 GSS.
 Zea mays
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 885)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citsek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 other GSSs: OGXND04TV
 Contract: Cathy Whitelaw

Query Match 28.0%; Score 384.2; DB 13; Length 659;

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1. .885
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0661B08"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 27.4%; Score 375.6; DB 29; Length 885;
Best Local Similarity 95.5%; Pred. No. 9.6e-35;
Matches 422; Conservative 1; Mismatches 5; Indels 14; Gaps 3;

QY 823 CAAAGTCAACTTCGACTTCGTCGCGCGCGTGCCTCGACCGCGCGCGGTTCCAGTG 882
DB 577 CAGGTACAACTTCGACTTCGTCGCGCGCGTGCCTCGACCGCGCGCGGTTCCAGTG 518
QY 883 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTGCCTCGAGTGAAGTGAAGCTAGAAAGA 942
DB 517 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTGCCTCGAGTGAAGTGAAGCTAGAAAGA 458
QY 943 GAAAGATCGCCCCCCCCCCCCCAACAAACATAACGAGAGAGAGAAAACCAACA 1002
DB 457 GAAAGATG-----CCCCCAACAAACATAACGAGAGAG-AAAACCGAACA 411
QY 1003 ATTAAGCAGTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGC 1062
DB 410 ATTAAGCAGTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGC 351
QY 1063 CTTGCTTTTCTCGAGCTAGCAGGAGCGTAGTTATTTATTTAGTACTTCTACTTTATTC 1122
DB 350 CTTGCTTTTCTCGAGCTAGCAGGAGCGTAGTTATTTATTTAGTACTTCTACTTTATTC 291
QY 1123 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 1182
DB 290 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 231
QY 1183 CTTAGTAGGAGTTAATCGTCTTAATTTATTTACTGTACAGCAGCTTGTGGCTTGCAAA 1242
DB 230 CTTAGTAGGAGTTAATCGTCTTAATTTATTTACTGTACAGCAGCTTGTG- TGGCTTGCAAA 172
QY 1243 GAAAGATCTGGTTTGTCTCAA 1264
DB 171 GAAAGATCTGGTTTGTCTCTAA 150

RESULT 10
CG223296 887 bp DNA linear GSS 22-AUG-2003
LOCUS OGWAU17TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0524C09,
genomic survey sequence.
DEFINITION CG223296
ACCESSION CG223296.1 GI:34123184
VERSION GSS.
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 887)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWAU17TH
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .887
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0524C09"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 27.4%; Score 375.6; DB 29; Length 887;
Best Local Similarity 95.5%; Pred. No. 9.6e-35;
Matches 422; Conservative 1; Mismatches 5; Indels 14; Gaps 3;

QY 823 CAAAGTCAACTTCGACTTCGTCGCGCGCGTGCCTCGACCGCGCGCGGTTCCAGTG 882
DB 133 CAGGTACAACTTCGACTTCGTCGCGCGCGTGCCTCGACCGCGCGCGGTTCCAGTG 192
QY 883 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTGCCTCGAGTGAAGTGAAGCTAGAAAGA 942
DB 193 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTGCCTCGAGTGAAGTGAAGCTAGAAAGA 252
QY 943 GAAAGATCGCCCCCCCCCCCCCAACAAACATAACGAGAGAGAGAAAACCAACA 1002
DB 253 GAAAGATG-----CCCCCAACAAACATAACGAGAGAG-AAAACCGAACA 299
QY 1003 ATTAAGCAGTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGC 1062
DB 300 ATTAAGCAGTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGC 359
QY 1063 CTTGCTTTTCTCGAGCTAGCAGGAGCGTAGTTATTTATTTAGTACTTCTACTTTATTC 1122
DB 360 CTTGCTTTTCTCGAGCTAGCAGGAGCGTAGTTATTTATTTAGTACTTCTACTTTATTC 419
QY 1123 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 1182
DB 420 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 479
QY 1183 CTTAGTAGGAGTTAATCGTCTTAATTTATTTACTGTACAGCAGCTTGTGGCTTGCAAA 1242
DB 480 CTTAGTAGGAGTTAATCGTCTTAATTTATTTACTGTACAGCAGCTTGTG- TGGCTTGCAAA 538
QY 1243 GAAAGATCTGGTTTGTCTCAA 1264
DB 539 GAAAGATCTGGTTTGTCTCTAA 560

RESULT 11
CAL06521 705 bp mRNA linear EST 23-SEP-2003
LOCUS SCQHR1013C07.9 HR1 Saccharum officinarum cDNA clone SCQHR1013C07
5', mRNA sequence.
DEFINITION CAL06521
ACCESSION CAL06521.1 GI:34959828
VERSION CAL06521
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.

```
REFERENCE
1 (bases 1 to 705)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
CONTACT: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 013 row: C column: 07
Seq primer: T7 Promoter Primer.
FEATURES
source
1..705
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCGFHR1013C07"
/lab_host="DH10B"
/clone_lib="HR1"
/notes="Organ: seedlings inoculated with Herbaspirillum
ruberisubalbicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Herbaspirillum
ruberisubalbicans]. cDNA was prepared from polyA+ mRNA
using Superscript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN
Query Match 27.3%; Score 375.2; DB 13; Length 705;
Best Local Similarity 82.0%; Pred. No. 1.2e-34;
Matches 557; Conservative 0; Mismatches 71; Indels 51; Gaps 9;
Qy 55 GCCTGTGGGAGGAAAAGAAAGAGAACCGCCCAAGCAAGCAAGCAGAGGCCAG 114
Db 31 GCAGCGGTGGGTGCGAGAGAAAAGAGAACCGCCCAAGCAAGCAAGCAGAGGCCAG 90
Qy 115 GSCC---CGGCGTTGCTCAGATGGGGAAGTACATGC---GCAAGTCAGCGGCGCGC 168
Db 91 GGCTCGCGCGGTTCCTGTCAGATGGGGAAGTACATGCAGCAAGTCAGCGGCGCGC 150
Qy 169 AGCGCGGAGGTGCGCGCGCTCGAGTTACGAGGTCTGCGCGCTCGCGAGCGAGTCCAG 228
Db 151 GGGCGAGAGGTGCGCGCGCTCGAGTCAAGCAGGTCTGCGCGCTCGCGAGCGAGTCCAG 210
Qy 229 GTCCGCGCGCGGAC---GGCGGTGTCGCAAGTTCGCCCGCGAGGAGCAAGAGGCGC- 284
Db 211 GTCCGCGCGCGGACCGCGGGGTGTGTGGCGAAGGTGCGCGAGGAGGAAGAGCGCT 270
Qy 285 -----CGCGGGGAGCCTGCTGTCGCCCGCTGAGCGCTGTGGGAGCGCGG 330
Db 271 GGCGCGCGCCAACTGTGGCGCGCGGGAGCTGTCTGCCGTGAGGGGACGGCGG 330
Qy 331 AAGTGTGTACATCCACCTCGGTAGCCGATGCTGTATGGAACCGCTCAGCGCGAGCC 390
Db 331 AAGTGTGTACATCCATCTCGGAGCGCGATGCTGTTCATGCGCACCGCTCAGCGCGAGCC 390
Qy 391 GTGGTTTACTCGGTTCGACCCCGGTGGAGGCTCTGATGGCGCTGAGGACACAGG 450
Db 391 GTGCG-----TTCCGACCCCGGAGGCTCTG-----TGAGGACACAGG 435
Qy 451 CGCGCGGTGCGGGCGGGGCTCTCGCGTTGTCTCCAGCAGCGCGTCTCGGTGAATCGG 510
Db 436 CGCGCGGTGCGGGCGGGGCTCTCGCGTTGTCTCCAGCAGCGCGTCTCGGTGACGTGG 495
```

```
Qy 511 CTTGGGGGCTCAGCGCGGAGCCACACTGCTCCCTCTACAGCTGCGAGGCTGGCGG 570
Db 496 ---GGGTCACTCAGCGCAAGAGCCAGCTCGCGCTC---CGACGCTCGGAGTTGGCGG 549
Qy 571 GGATCAGCTCTCTGTGATGTCCTCGGCGGCGAGCAACTCCGGGAGCGGCCAGACCGCGA 630
Db 550 GGATCAGCTCTCTGTGATGTCCTCGGCGGCGAGCAACTCCGGGAGCGGCCAGACCGCGA 603
Qy 631 GAGCGGAGAGAGCAGCGCATCTGAGCGCGGCGGCGAGCTCAGCGATCTGGAGTCCGA 690
Db 604 GAGGAGAGAGAGCAGCGCATCTGAGCGCGGCGGCGAGCTCAGCGATCTGGAGTCCGA 663
Qy 691 TCTGCGCGGCGCAAGACT 709
Db 664 TCGGCGGCGGAGCAAGATT 682
```

```
RESULT 12
CA104638/c
LOCUS
DEFINITION SCJFHR1034B05.b HR1 Saccharum officinarum cDNA clone SCJFHR1034B05
3', mRNA sequence.
ACCESSION CA104638
VERSION CA104638.1 GI:34957945
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
REFERENCE 1 (bases 1 to 655)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 034 row: B column: 05
Seq primer: SP6 Promoter primer.
FEATURES
source
1..655
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFHR1034B05"
/lab_host="DH10B"
/clone_lib="HR1"
/notes="Organ: seedlings inoculated with Herbaspirillum
ruberisubalbicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Herbaspirillum
ruberisubalbicans]. cDNA was prepared from polyA+ mRNA
using Superscript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN
Query Match 27.3%; Score 374.4; DB 13; Length 655;
Best Local Similarity 82.6%; Pred. No. 1.5e-34;
Matches 542; Conservative 0; Mismatches 66; Indels 48; Gaps 8;
Qy 150 GCAAGTGCAGGGCGCGCGAGGTCGCGCGGCGGAGTTCAGGATTACGAGGTCGTCG 209
```

Db 629 GCAAGTCCAGGGGACCGCGGGGAGAGGTTCCCGGTTTCGAGCTCACGCAGATCGTCG 570
QY 210 GCGTCCGGAGAGTCCAGTCCGCGCGGCGGACCGCGG---TGTGCGAAAGTCCGCC 266
Db 569 GCGTCCGGAGAGTCCAGATCCGCGCGGCGGACCGCGGGTGTGTGTGCGAAAGATCGCAC 510
QY 267 CGAGGAGGAGAGGGGCGC-----CGGCGGGGGAGCCCTGCTGCCCGCGGTGA 311
Db 509 CGAGGAGGAGAGGCGCTGGCGCGCTCCAAAGTGGCGGGGAGGAGCTGTGCGCGTCG 450
QY 312 GCGCTGGTGGGAGCGCGGAAAGTGTACATCCACCTCGGTAGCCGCATGCTGTTCATGG 371
Db 449 GCGTCCGAGGGGAGCGCGGAAAGTGTACATCCATCTGCGCAGCGCATGCTGTTCATGG 390
QY 372 CACCGCTCAGCGCGAGCGCTGCGGTTGACCTCGGTTCCGACCCCGGTGGAGGCTGCTGATG 431
Db 389 CACCGCTTACGCGCAGCTG-----TCGCTTCCGACCGCGGGAGGTTGCT---- 343
QY 432 GCGCTCAGGACAGCAGGCGCGCGCTCGCGGCGGGGCTCTCGCGTTGCTCCAGACCGG 491
Db 342 --GGTCCAGACAGCAGGCGCGCGCTCGCGCGGGGCTCTCGCGTTGCTCCAGACCGG 285
QY 492 CFTGCTCGGTGAACCTGGGCTTGGGGGTTCAGCGCGGAGCCACACTCGCGCTCTCTACG 551
Db 284 CFTGCTCGGTGACGTCG---GGGTTCATCAGCGCAGGAGCCACGCTCGCGCTC---CG 231
QY 552 ACCTCAGAGGCTGGCGGGATCAGTCTGCTGGTGGATGCTCGCGCGGCGAGCAACTCCG 611
Db 230 ACCTCAGAGGTTGGCGGGATCAGCCCGGATG-----TCTCGCGAGCAACTCCG 177
QY 612 GAGCGCGCCAGACCGCAGAGGCGAGAGACGCGCATCAGCGCGGGCGCACGGCGAGC 671
Db 176 GAGCGCGCCAGACCGCAGAGGAGAGAGACGCGCATCAGCGCGGGCGCACGGCGAGC 117
QY 672 TCAGCATCTGAGTCTGATCTGGCGGGGACCAAG---ACTGCGCGCTGCTACCGCGCGG 728
Db 116 TCAGCATCTGAGTCTGATCTGGCGGGGACCAAGACTACTGCGCGCTGCTGCGCGCGG 57
QY 729 CACGCGCGCTGGGAGCTGATCGTCCGCGCAGCAGCAGATCCAGGATCTTTC 784
Db 56 CACGCTCGGCTGGCAGCGATCGTGCCACCAAGAGATCGAGGAGTTCTTC 1

RESULT 13
CC390573 790 bp DNA linear GSS 19-MAY-2003
LOCUS FUHBT67D.ZM.0.6.1.0.KB.Zea.mays.genomic.clone.ZM8Bta415M08,
DEFINITION genomic survey sequence.

ACCESSION CC390573
VERSION CC390573.1 GI:30870663
KEYWORDS GSS.
SOURCE Zea mays

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 790)
Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Benetzen, J.

Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHBT67B
Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

Fax: 301-838-0208
Email: whitelaw@tigr.org

Seq primer: TF
Class: sheared ends.

Location/Qualifiers
1..790

FEATURES
source

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8Bta415M08"
/clone_lib="ZM.0.6.1.0.KB"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match. 26.3%; Score 361.2; DB 28; Length 790;
Best Local Similarity 95.0%; Pred. No. 4.8e-33;
Matches 420; Conservative 1; Mismatches 4; Indels 17; Gaps 4;
QY 823 CAAGTCAACTTCGACTTCGTCGCGCGGTGCCCCCTCGACGCGCGCGCGGTTCGAGTG 882
Db 510 CAGGTCAACTTCGACTTCGTCGCGGTGCGGTGCCCCCTCGACGCGCGCGGTTCGAGTG 451
QY 883 GCGCGCGGTGCTCAGCATCTGAAGCGAGCGGTGCGGTGCAAGGTGAAGTAGAAGA 942
Db 450 GCGCGCGGTGCTCAGCATCTGAAGCGAGCGGTGCGGTGCAAGGTGAAGTAGAAGA 391
QY 943 GAAAGATGCCCCCCCCCCCCCAACAAACATAACCGAGAGAGAGAGAGAGAGAGAG 1002
Db 390 GAAAGATG-----CCCCCCCCCAACAAACATAACCGAGAGAGAGAGAGAGAGAG 341
QY 1003 ATTAAGCAGCTTTATATAGCTTAAGCTAACCCACCATTCATCTGTCGCAATGCATGC 1062
Db 340 ATTAAGCAGCTTTATATAGCTTAAGCTAACCCACCATTCATCTGTCGCAATGCATGC 285
QY 1063 CTTGCTTTCTCGGAGCTAGCAGGAGCGTAGTATTATTTAGTACTACTTTACTTATTC 1122
Db 284 CTTGCTTTCTCGGAGCTAGCAGGAGCGTAGTATTATTTAGTACTACTTTACTTATTC 225
QY 1123 AGAGGTTATCTTACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 1182
Db 224 AGAGGTTATCTTACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 167
QY 1183 CTTAGATGAGTTTAAATCGCTTTAATTTATTAATCTGACAGCAGCTTGTGGCTTGCAAA 1242
Db 166 CTTAGATGAGTTTAAATCGCTTTAATTTATTAATCTGACAGCAGCTTGTGGCTTGCAAA 108
QY 1243 GAAAGATCTGTTTGTCTCAA 1264
Db 107 GAAAGATCTGTTTGTCTCTAA 86

RESULT 14
CC218361/c

LOCUS CG218361

DEFINITION

genomic survey sequence.

ACCESSION

CG218361

VERSION

CG218361.1 GI:34118249

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 835)

Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Other GSSs: OGIDJ10TV

Unpublished (2002)

Consortium for Maize Genomics

Other GSSs: OGIDJ10TV

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

Location/Qualifiers
 1..835
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0736B20"
 /clone_lib="ZM_0.7_1.5_KB"
 /note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 26.3%; Score 361.2; DB 29; Length 835;
 Best Local Similarity 95.0%; Pred. No. 4.7e-33;
 Matches 420; Conservative 1; Mismatches 4; Indels 17; Gaps 4;

QY 823 CAAGTACAACTTCGACCTTCGCGCGGTGCCCCCTCGACCGCGCGCGGTTCGAGTG 882
 DB 585 CAGGTACAACTTCGACCTTCGCGCGGTGCCCCCTCGACCGCGCGCGGTTCGAGTG 526
 QY 883 GGCGCGGTGTCGACATCTGAAGCGAGCGTGGCGTCCGCTCAAGGTGAAGCTAGAAAGA 942
 DB 525 GGCGCGGTGTCGACATCTGAAGCGAGCGTGGCGTCCGCTCAAGGTGAAGCTAGAAAGA 466
 QY 943 GAAAGATGCCCCCCCCCCCCCCCCCAACAAACATACGAGAGAGAAACCAACA 1002
 DB 465 GAAAGATG-----CCCCCCCCCAACAAACATACGAGAGAGAAACCAACA 416
 QY 1003 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACTTCATCTCGTCCAAATGCAATGC 1062
 DB 415 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACTTCATCTCGTCCAAATGCAATGC 360
 QY 1063 CTTCGTTTCTCGAGCTAGCAGAGCGTAGTTATTTATTTAGTACTACTTTACTTATTC 1122
 DB 359 CTTCGTTTCTCGAGCTAGCAGAGCGTAGTTATTTATTTAGTACTACTTTACTTATTC 300
 QY 1123 AGAGTTATCTGACCCGATAGATCAATCGCTTACGTGTAAATTTCTCTCATGCACT 1182
 DB 299 AGAGTTATCTGACCCGATAGATCAATCGCTTACGTGTAAATTTCTCTCATGCACT -T 242
 QY 1183 CTTAGATCGAGTTAATCGCTTAAATTTATTTACTCTACAGCAGCTTGTGTGGCTTGCAAA 1242
 DB 241 CTTAGATCGAGTTAATCGCTTAAATTTATTTACTCTACAGCAGCTTGC-TGGCTTGCAAA 183
 QY 1243 GAAAGATCGTTTGTCTCAAA 1264
 DB 182 GAAAGATCGTTTGTCTCTAA 161

RESULT 15
 CG265472/c
 LOCUS CG265472 825 bp DNA linear GSS 25-AUG-2003
 DEFINITION OG2CM60TH ZM 0.7_1.5 KB Zea mays genomic clone ZMMBMA0764123,
 genomic survey sequence.

ACCESSION CG265472

VERSION CG265472.1 GI:34177613

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 825)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.

TITLE Consortium for Maize Genomics

JOURNAL Unpublished (2002)

COMMENT Other_GSSs: OG2CM60TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..825
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0764123"
 /clone_lib="ZM 0.7_1.5_KB"
 /note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 26.0%; Score 356.6; DB 29; Length 825;
 Best Local Similarity 95.4%; Pred. No. 1.6e-32;
 Matches 392; Conservative 1; Mismatches 5; Indels 13; Gaps 2;

QY 823 CAAGTACAACTTCGACCTTCGCGCGGTGCCCCCTCGACCGCGCGCGGTTCGAGTG 882
 DB 399 CAGGTACAACTTCGACCTTCGCGCGGTGCCCCCTCGACCGCGCGCGGTTCGAGTG 340
 QY 883 GGCGCGGTGTCGACATCTGAAGCGAGCGTGGCGTCCGCTCAAGGTGAAGCTAGAAAGA 942
 DB 339 GGCGCGGTGTCGACATCTGAAGCGAGCGTGGCGTCCGCTCAAGGTGAAGCTAGAAAGA 280
 QY 943 GAAAGATGCCCCCCCCCCCCCCCCCAACAAACATACGAGAGAGAAACCAACA 1002
 DB 279 GAAAGATG-----CCCCCAACAAACATACGAGAGAGAG-AAAACCGAACA 233
 QY 1003 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACTTCATCTCGTCCAAATGCAATGC 1062
 DB 232 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACTTCATCTCGTCCAAATGCAATGC 173
 QY 1063 CTTCGTTTCTCTCGAGCTAGCAGAGCGTAGTTATTTATTTAGTACTACTTTACTTATTC 1122
 DB 172 CTTCGTTTCTCTCGAGCTAGCAGAGCGTAGTTATTTATTTAGTACTACTTTACTTATTC 113
 QY 1123 AGAGTTATCTGACCCGATAGATCAATCGCTTACGTGTAAATTTCTCTCATGCACT 1182
 DB 112 AGAGTTATCTGACCCGATAGATCAATCGCTTACGTGTAAATTTCTCTCATGCACT 53
 QY 1183 CTTAGATCGAGTTAATCGCTTAAATTTATTTACTCTACAGCAGCTTGTGTG 1233
 DB 52 CTTAGATCGAGTTAATCGCTTAAATTTATTTACTCTACAGCAGCTTGTGTG 2

Search completed: October 2, 2004, 04:18:56

Job time : 17939 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 23:19:55 ; Search time 704 Seconds
(without alignments)
9879.895 Million cell updates/sec

Title: US-09-993-808B-1

Perfect score: 1372

Sequence: 1 cccagcgctccgagcgaag.....agtgtcacctaaattcattc 1372

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1371.6	100.0	1372	9	US-09-993-308-1
2	1371.6	100.0	1372	11	US-09-993-808B-1
3	990.8	72.2	1111	13	US-10-425-114-34947
C 4	869.6	63.4	985	13	US-10-425-114-18379
C 5	360.4	26.3	834	13	US-10-425-114-18085
6	351	25.6	841	9	US-09-993-308-5
7	351	25.6	841	11	US-09-993-808B-5
C 8	301	21.9	1671	17	US-10-333-963-92632
9	299	21.8	1242	16	US-10-333-006-9
10	272.4	19.9	510	17	US-10-767-701-24179
11	219	16.0	568	16	US-10-333-006-5
12	114.8	8.4	2015	13	US-10-653-595-104
13	114.8	8.4	2015	13	US-09-397-945-104
14	110.8	8.1	3313	9	US-09-764-846-83

15	110.8	8.1	3313	15	US-10-091-483-83	Sequence 83, Appl
16	110.2	8.0	1557	9	US-09-925-297-262	Sequence 262, App
17	110.2	8.0	1557	15	US-10-023-896-42	Sequence 42, Appl
18	109.6	8.0	701	15	US-10-106-698-519	Sequence 519, Appl
19	108.4	7.9	149	10	US-09-764-891-2570	Sequence 2570, Ap
20	108.4	7.9	2116	9	US-09-925-297-261	Sequence 261, App
21	108.4	7.9	2116	15	US-10-023-896-31	Sequence 31, Appl
22	107.4	7.8	577	9	US-09-925-301-628	Sequence 628, App
23	107.4	7.8	1837	9	US-09-925-297-146	Sequence 146, App
C 24	107.4	7.8	2140	9	US-09-764-853-250	Sequence 250, Appl
25	107.2	7.8	2322	15	US-10-194-125-3	Sequence 3, Appli
C 26	106.8	7.8	845	16	US-10-264-049-133	Sequence 133, App
27	106.6	7.8	858	15	US-10-106-698-201	Sequence 201, App
28	106.4	7.8	584	9	US-09-925-299-312	Sequence 312, App
29	106.4	7.8	584	10	US-09-925-299-312	Sequence 313, App
30	106.4	7.8	763	9	US-09-925-299-313	Sequence 29, Appl
31	106.4	7.8	763	10	US-09-925-299-313	Sequence 29, Appl
32	106.4	7.8	824	15	US-10-144-929-29	Sequence 1, Appl
33	106.4	7.8	824	16	US-10-144-929-29	Sequence 319, App
34	105.6	7.7	2402	15	US-10-194-125-1	Sequence 319, App
35	105.4	7.7	418	9	US-09-925-299-319	Sequence 319, App
36	105.4	7.7	418	10	US-09-925-299-319	Sequence 2584, Ap
37	105.4	7.7	627	10	US-09-764-891-2584	Sequence 1342, Ap
38	105.4	7.7	1280	15	US-10-106-698-1342	Sequence 75, Appl
39	105.4	7.7	1280	16	US-10-264-049-75	Sequence 93, Appl
40	105.2	7.7	2454	15	US-10-050-704-93	Sequence 330, App
41	105.2	7.7	2454	17	US-10-798-512-93	Sequence 315, App
42	104.4	7.6	453	9	US-09-925-299-330	
43	104.4	7.6	453	10	US-09-925-299-330	
44	104.4	7.6	540	9	US-09-925-299-315	
45	104.4	7.6	540	10	US-09-925-299-315	

ALIGNMENTS

RESULT 1

US-09-993-308-1
; Sequence 1, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dalkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)..(904)
; OTHER INFORMATION:
; US-09-993-308-1

Query Match 100.0%; Score 1371.6; DB 9; Length 1372;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CCACGCGTCCGAGCGAAGCGGTGCGAGGAGCGCGCGAGCGGTTCGCGCTT	60
Db	1	CCACGCGTCCGAGCGAAGCGGTGCGAGGAGCGCGCGAGCGGTTCGCGCTT	60
Qy	61	GGGAGGAAAGAGAGAGAGACCGCCGCGAGACGAGAGCGAGCGCGC	120

Db	61	GGGAGAGAAAAAGAGAAAAGAGAAACGGGCTAAGACAAGCAAGCCGAGAGAGCCAGGCGCGC	120
Qy	121	GGGCTTCGCTCAGATGGGGAAGTACATGCGCAAGTGCAGGGGCGCGCAGGCGCGAGGT	180
Db	121	GGGCTTCGCTCAGATGGGGAAGTACATGCGCAAGTGCAGGGGCGCGCAGGCGCGAGGT	180
Qy	181	CGCGCGCTCGAGGTATACGAGTCTGTCGCGCTCCGACGAGTCCAGGTTCGCGCGCGC	240
Db	181	CGCGCGCTCGAGGTATACGAGTCTGTCGCGCTCCGACGAGTCCAGGTTCGCGCGCGC	240
Qy	241	GACCGCGCTTCGCGAAGTTCGCCCCGAGAGAGAGAGGCGCCCGCGGGGAGCCTGC	300
Db	241	GACCGCGCTTCGCGAAGTTCGCCCCGAGAGAGAGAGGCGCCCGCGGGGAGCCTGC	300
Qy	301	TGCGCGCTGAGCGCTGGTGGGACGCGGAAGTCTATCATCACTGCGTAGCCGAT	360
Db	301	TGCGCGCTGAGCGCTGGTGGGACGCGGAAGTCTATCATCACTGCGTAGCCGAT	360
Qy	361	GCTGTTTATGCAACCGCTCAGCGCAGCGTTCGTTGACTCGGTTCCGACCCCGGTGGA	420
Db	361	GCTGTTTATGCAACCGCTCAGCGCAGCGTTCGTTGACTCGGTTCCGACCCCGGTGGA	420
Qy	421	GGCTGCTGATGGCGCTGCAGACACAGGCGCGCGCTCCGCGCGGGCTTCGCGTTG	480
Db	421	GGCTGCTGATGGCGCTGCAGACACAGGCGCGCGCTCCGCGCGGGCTTCGCGTTG	480
Qy	481	CTCCAGCAACGCGCTGCTCGGTGAACCTTGGGCTTGGGGGTTCAGCGCGGAGCACACCTG	540
Db	481	CTCCAGCAACGCGCTGCTCGGTGAACCTTGGGCTTGGGGGTTCAGCGCGGAGCACACCTG	540
Qy	541	CCGCTTCCTACGACCTGCAGAGGCTGGCGGGATCACGTCTGGTGAATGTCGCGCGC	600
Db	541	CCGCTTCCTACGACCTGCAGAGGCTGGCGGGATCACGTCTGGTGAATGTCGCGCGC	600
Qy	601	GAGCAACTCCGGAGCGCGCCAGACCCGCGAGGCGAGACGACGCCATCGAGCGCGGC	660
Db	601	GAGCAACTCCGGAGCGCGCCAGACCCGCGAGGCGAGACGACGCCATCGAGCGCGGC	660
Qy	661	GCACGGGAGCTCAGCGATCTGGAGTGGATCTCGCGCGGCGACAGACTGGCCGCTGCT	720
Db	661	GCACGGGAGCTCAGCGATCTGGAGTGGATCTCGCGCGGCGACAGACTGGCCGCTGCT	720
Qy	721	ACCGCGCAACGCGCGCTCGGAGCTGATCGTCCGCGCAGACACAGAGTCCAGAGTT	780
Db	721	ACCGCGCAACGCGCGCTCGGAGCTGATCGTCCGCGCAGACACAGAGTCCAGAGTT	780
Qy	781	CTTCGCGCGCGCGAGCGCGCCAGCGCTTCGTTCCAAGTACAACTTCGACTT	840
Db	781	CTTCGCGCGCGCGAGCGCGCCAGCGCTTCGTTCCAAGTACAACTTCGACTT	840
Qy	841	CGTCGCGCGTGCCTCTGACGCGCGCGCGCTTCAGTGGGCGCGGTGCTCAGAT	900
Db	841	CGTCGCGCGTGCCTCTGACGCGCGCGCGCTTCAGTGGGCGCGGTGCTCAGAT	900
Qy	901	CTGAAGCAGCTGCTCGCTCCGCTGCAAGTGAAGCTAGAAAGAGAAAAGATGCCCGCCCC	960
Db	901	CTGAAGCAGCTGCTCGCTCCGCTGCAAGTGAAGCTAGAAAGAGAAAAGATGCCCGCCCC	960
Qy	961	CCCCCCCCCAACAAATACGGAAGAGAAAAACCAACAAATAAGAGCTTTATATA	1020
Db	961	CCCCCCCCCAACAAATACGGAAGAGAAAAACCAACAAATAAGAGCTTTATATA	1020
Qy	1021	GCCTAAGCTAACCAACCAATCATCTGCTCAATGATGCTGCTTCTCTCGGAGC	1080
Db	1021	GCCTAAGCTAACCAACCAATCATCTGCTCAATGATGCTGCTTCTCTCGGAGC	1080
Qy	1081	TAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTTTATTCAGAGGTATCTTGACCCC	1140
Db	1081	TAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTTTATTCAGAGGTATCTTGACCCC	1140
Qy	1141	GATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCACTCTTAGTAGGAGTTTAACT	1200

Db	1141	GATAGATCAANCGCGTTACTGTGTAAATTTCTCTCATGCATCTCTTAGATGGAGTTTAAATC	1201
Qy	1201	GTCTTAATTATTACTCTACAGCAGCTTGTTGGCTTGCAAAGAAAGATCTGGTTTGTCT	1260
Db	1201	GTCTTAATTTTACTGTGTACAGCAGCTTGTTGGCTTGCAAAGAAAGATCTGGTTTGTCT	1260
Qy	1261	CAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTTAGAGGATCCAAGCTTTACG	1320
Db	1261	CAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTTAGAGGATCCAAGCTTTACG	1320
Qy	1321	TACGCGTGCATGGCAGCTCATAGCTCTTCTATAGTGTACCTAAATTCATTC	1372
Db	1321	TACGCGTGCATGGCAGCTCATAGCTCTTCTATAGTGTACCTAAATTCATTC	1372

RESULT 2

```

RESULT 2
US-09-993-808B-1
; Sequence 1, Application US/09993808B
; Publication No. US2004000343A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)...(902)
US-09-993-808B-1

```

Query Match	100.0%; Score 1371.6; DB 11; Length 1372;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1	CCCACGCGTCCGGACGCAAGCGGTGCAGGCAGCAGCGCCGCGCAGAGCGTTGTGGCCCTGT 60
DB	
QY 1	CCCACGCGTCCGGACGCAAGCGGTGCAGGCAGCAGCGCCGCGCAGAGCGTTGTGGCCCTGT 60
DB	
QY 61	GGGAGAGGAAAAAGAAAGAGGAACCGGCCAAAGCAAGCAAGCGAGAGGCCAGAGGCCGC 120
DB	
QY 61	GGGAGAGGAAAAAGAAAGAGGAACCGGCCAAAGCAAGCAAGCGAGAGGCCAGAGGCCGC 120
DB	
QY 121	GGCGTTGCGTCAGATGGCGGAGTACATCGCAAGTGCAGGCGCGCCGACGCGCGCGAGGT 180
DB	
QY 121	GGCGTTGCGTCAGATGGCGGAGTACATCGCAAGTGCAGGCGCGCCGACGCGCGCGAGGT 180
DB	
QY 181	CGCCGCCGTCGAGGTTACGCAAGTTCGTCGGCGTCGGACGAGGTCCAGTCCGCGCGCGC 240
DB	
QY 181	CGCCGCCGTCGAGGTTACGCAAGTTCGTCGGCGTCGGACGAGGTCCAGTCCGCGCGCGC 240
DB	
QY 241	GCACGGCGGTTCGCGGAAGTTCGCCCGAGAGGAGAGGGCGCGCGGGGGGAGACCTTCG 300
DB	
QY 241	GCACGGCGGTTCGCGGAAGTTCGCCCGAGAGGAGAGGGCGCGCGGGGGGAGACCTTCG 300
DB	
QY 301	TGCCCCGGTGACGCGTGTGTGGGACGGCGGAAGTGTCTACATCCACCTTCGCTAGCCGCAT 360
DB	
QY 301	TGCCCCGGTGACGCGTGTGTGGGACGGCGGAAGTGTCTACATCCACCTTCGCTAGCCGCAT 360
DB	
QY 361	GCTGTTTCATGGCAACCGCTCAGCCGACGCGTTCGCTCGGTTCGACATCCCGCGTGGGA 420
DB	
QY 361	GCTGTTTCATGGCAACCGCTCAGCCGACGCGTTCGCTCGGTTCGACATCCCGCGTGGGA 420
DB	

```
421 GGCTGCTGATGGCTGACAGGACAGCAGCGCGCGCTCGCGCGCGCTCGCGTTG 480
421 GGCTGCTGATGGCTGACAGGACAGCAGCGCGCGCTCGCGCGCGCTCGCGTTG 480
481 CTCACGACGGCGTCTCGGTGAACCTTGGGCTTGGGGGTTCAGCGCGGAGCCACCTG 540
481 CTCACGACGGCGTCTCGGTGAACCTTGGGCTTGGGGGTTCAGCGCGGAGCCACCTG 540
541 CGCTCTTACGAGCTGACAGGCTGCGGGATCACGCTCTGCTGCTGCTGCTGCTG 600
541 CGCTCTTACGAGCTGACAGGCTGCGGGATCACGCTCTGCTGCTGCTGCTGCTG 600
601 GAGCAACTCCGCGAGCGCGCCAGACCGCAGAGGCGAGAGACGACGCCATCGAGCGCGC 660
601 GAGCAACTCCGCGAGCGCGCCAGACCGCAGAGGCGAGAGACGACGCCATCGAGCGCGC 660
661 GACGCGGAGCTCAGCGATCTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 720
661 GACGCGGAGCTCAGCGATCTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 720
721 ACCGGGCGCAACCGCGCTCGGAGCTGATCGTCCCGCCAGCAGCAGCAGATCCAGAGTT 780
721 ACCGGGCGCAACCGCGCTCGGAGCTGATCGTCCCGCCAGCAGCAGCAGATCCAGAGTT 780
781 CTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
781 CTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
841 CTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
841 CTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
901 CTTGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
901 CTTGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
961 CTTGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
961 CTTGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
1021 GCTTAAGCTTAACCAACCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
1021 GCTTAAGCTTAACCAACCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
1081 TAGCGAGCGAGTATTTATTTAGTACCTTATTTAGTACCTTATTTAGTACCTTAT 1140
1081 TAGCGAGCGAGTATTTATTTAGTACCTTATTTAGTACCTTATTTAGTACCTTAT 1140
1141 GATAGATCAATCCGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
1141 GATAGATCAATCCGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
1201 GTCTTAATTTATTTAGTACCTTATTTAGTACCTTATTTAGTACCTTATTTAGT 1260
1201 GTCTTAATTTATTTAGTACCTTATTTAGTACCTTATTTAGTACCTTATTTAGT 1260
1261 CAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
1261 CAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
1321 TACGCGTGCATCGAGCTCATAGCTTCTTATAGTGTACCTTAAATTCATTC 1372
1321 TACGCGTGCATCGAGCTCATAGCTTCTTATAGTGTACCTTAAATTCATTC 1372
```

RESULT 3

US-10-425-114-34947

; Sequence 34947, Application US/10425114

; Publication No. US2004003488A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

```
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34947
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17355F05_FLI
US-10-425-114-34947
```

```
Query Match 72.2%; Score 990.8; DB 13; Length 1111;
Best Local Similarity 96.0%; Pred. No. 1.3e-259;
Matches 1089; Conservative 1; Mismatches 13; Indels 31; Gaps 6;

QY 103 GCGAGAGGCGAGGCGCGCGCGTTCGCTCAGATGGGGAAGTACATCGCAAGTGCAGGGG 162
DB 1 GCGAGAGGCGAGGCGCGCGCGTTCGCTCAGATGGGGAAGTACATCGCAAGTGCAGGGG 60

QY 163 CCGCGCGAGCGCGGAGTTCGCGCGCGTTCGAGGTTACGAGGTCGTCGGGTCGGGACGAG 222
DB 61 GCGCGAGCGCGGAGTTCGCGCGCGTTCGAGGTTACGAGGTCGTCGGGTCGGGACGAG 120

QY 223 GTCCAGGTCGCGCGCGCGCGCGGCGGTCGCGGAGTTCGCGGAGTTCGCGGAGTTCG 279
DB 121 GTCCAGGTCGCGCGCGCGCGCGGCGGTCGCGGAGTTCGCGGAGTTCGCGGAGTTCG 180

QY 280 GCGCGCGCGCGGAGGCGGTCGTCGCGCGTTCGAGCGTTCGTCGGGACGCGGAGTTCG 339
DB 181 GCGCGCGCGCGGAGGCGGTCGTCGCGCGTTCGAGCGTTCGTCGGGACGCGGAGTTCG 240

QY 340 CATCCACCTGCTAGCGCGCATGCTTCATGACCGCTCAGCGCGTTCGCGGAGTTCG 399
DB 241 CATCCACCTGCTAGCGCGCATGCTTCATGACCGCTCAGCGCGTTCGCGGAGTTCG 295

QY 400 CTCGGTTCGCGCGCGCGCGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 459
DB 296 -TCGGTTCGCGCGCGCGCGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 354

QY 460 CCGCGCGCGGCTCTCGCGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 519
DB 355 CCGCGCGCGGCTCTCGCGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 408

QY 520 TCAGCGCGGAGCGCACCTGCGCTCCTACGACGCTGCGAGGCTGCGGCGGATCACGT 579
DB 409 TCAGCGCGGAGCGCACCTGCGCTCCTACGACGCTGCGAGGCTGCGGCGGATCACGT 468

QY 580 CTTGTTGATGTCGTCGCGCGCGAGCAACTCCGCGGAGCGCGCGCGCGCGAGCGAG 639
DB 469 CTTGTTGATGTCGTCGCGCGCGAGCAACTCCGCGGAGCGCGCGCGCGCGAGCGAG 528

QY 640 GACGAGCGCATCAGCGCGCGCGAGCGCGAGTTCGAGGATTCGAGGATTCGCGCGG 699
DB 529 GACGAGCGCATCAGCGCGCGCGAGCGCGAGTTCGAGGATTCGAGGATTCGCGCGG 588

QY 700 GCACAGACTGCGCGCTGCTACCGCGCGCAACCGCGCGGCTGCGGAGTTCGTCGCG 759
DB 589 GCACAGACTGCGCGCTGCTACCGCGCGCAACCGCGCGGCTGCGGAGTTCGTCGCG 648

QY 760 AGCACACGAGATCCAGGAGTTCCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTC 819
DB 649 AGCACACGAGATCCAGGAGTTCCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTC 708

QY 820 TTCCAAGTACAACTTCGACTTCGTCGCGCGCTGCGCGCTTCGACGCGCGCGCGTTCGA 879
DB 709 TTCCAAGTACAACTTCGACTTCGTCGCGCGCTGCGCGCTTCGACGCGCGCGCGTTCGA 768
```


Qy 748 GATCGTCCGCCGACACAGAGATCCAGGAGTTCTTCGCCGCCGC 792

Db 797 GATCGTCCGCCGACACAGAGATCCAGGAAATTTCTTCGCCGCCGC 841

RESULT 7

```

US-09-993-808B-5
; Sequence 5, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lower, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993, 808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 841
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(841)
; OTHER INFORMATION: r= g or a
US-09-993-808B-5

```

Query Match.	25.6%;	Score 351;	DB 11;	Length 841;
Best Local Similarity	73.9%;	Pred. No. 4.1e-85;		
Matches 610;	Conservative 1;	Mismatches 136;	Indels 78;	Gaps 10;

Qy	22	GGCTGCAGGCAGACGCGCGCAGGCGTGTGTGCTTGGAGAGCAAAAGCAAGA 81
Db		
Db	41	GCTTGCAGGCAGCGCCGCGAGGCGTGGAGTGGCCAGTGGGAGTGGAGTGAAGA 100
Qy	82	GGAAACCGCCCAAGAACAAGCAAGCAGAGCGCAGG-----GCCCGCGCGTTGCCGTCAAGT 135
Db		
Db	101	GGAAACCGCCCAAGAGCAAGCAGAGGAAGGCAGTCTGCGCGCGGCTTCGTAAAGT 160
Qy	136	GGGGAATATCATGCGAAGTCAGGGGCGCCGACAGGCGCGAGAGTCGCGCGTCGAGGT 195
Db		
Db	161	GGGGAATATCATGCGCAAGCAGCAGGGGCGCCGCGCGGAGGGGTGCGCAGTCGAGGT 220
Qy	196	TAGCAGGTCTGCGGCTCCGGACGAGTCCAGTCCGCGCGCGCGAC---CGCGCGTGT 252
Db		
Db	221	CTCCAGGTCTGCGGCTCCGGAAGAGTCCAGTCCGCGCGCGACCGCGCGGTGT 280
Qy	253	CGCAAGGTGCGCCCGCAGGAGGA-----GGCGGCC 285
Db		
Db	281	CGCAAGGTGCTCGCGGAGGAGAGAGGCGCTGCTCCCGCCGCGACGTGAAC 340
Qy	286	GGCGGGAGCGTCTGCTCGCGCGTGAAGCGTGTGGGACGGCGAA---GCTGCTACAT 342
Db		
Db	341	GTGCGGGAGCCTGGTGCCTGGCGTGGTGTGGGAACGGCGGAAGTCTCTACAT 400
Qy	343	CCACCTGCTAGCCGATGCTGTTTCATGGACCGCCTCAGCGCAGCGCTCGTTGATC 402
Db		
Db	401	CCACCTCGGAGCGCATGCTGTTTCATGGACACCTCAGCAGCAACCGTCGG-----C 454
Qy	403	GGTTCACACCCCGTGAAGGCTGCTGATGGCGTGCAGACAGCAGCGCGCGCGCTCGC 462
Db		
Db	455	GGCTCTAGCCCGTGGAGGCTGTGTGTGGGCACACAGGCGGGGTGTGGCGCTCGC 514
Qy	463	GGCGGGCTCTCGGTTGCTCCAGCAGCGGCTCGTCGGTGAACCTTGGGGGTGCA 522
Db		
Db	515	GGCTGGCTCTCGGTTGCTCCAGCAGCGGCTCGTCGTGACGT----- 559

Qy	523	CGCGGGAGGCCA	CACCTTGC	CGCTCTTACGA	CGTTCG	GAGGCTGC	CGGGATCA	CGTCT	582		
Db	560	---	CGGGGGCAC	CCCTCG	CTCCGCTCG	AGCTCG	CGCTCG	GAGTTG	CGGGATCA	CGT	616
Qy	583	GGTGGATG	TTCTCG	CGCGGACCA	CTCG	GGAGCGGCC	CAGACCG	CGAGAGCG	GAGAGAC	642	
Db	617	CCCGGATG	TGTCA	CCGCGGACCA	CTCG	GGAGCGT	CCCGAC	CGCGAGAGAG	GAGAGAC	676	
Qy	643	GAGCCCA	---	TCGAGCGGGCGCA	---	CGGCGAGCT	CAGCAT	TTGGAGT	CGGATCTGC	696	
Db	677	GAGCCCAT	CGTCG	AGCGGGCGGCGAC	GGCGGCGAGT	CAGCGAT	TTGGAGT	CGGATCTGC	736		
Qy	697	GG	---	GGCACAGACTGC	-----	CGTTCCT	ACCGCGCGCAC	CGCGGT	TCCGAGCT	747	
Db	737	GGGCGCGG	AGAACTGG	TCTCTG	TCTG	TCTCG	CGCGGACAA	CAT	CGGCTCCGAGCT	796	
Qy	748	GATCGTCC	CGCCAGCA	CACGAGAT	TCAGGAGT	TTTCG	CCCGCGC	792			
Db	797	GATCGTCC	CGCCAGCA	CAGGAGAT	TCAGGAGT	TTTCG	CGCGCGC	841			

RESULTS

```

US-10-437-963-92632/c
; Sequence 92632, Application US/10437963
; Publication No. US20040123343A1
;
; GENERAL INFORMATION:
;
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
;
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
;
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;
; FILE REFERENCE: 38-21(53221)B
;
; CURRENT APPLICATION NUMBER: US/10/437,963
;
; CURRENT FILING DATE: 2003-05-14
;
; NUMBER OF SEQ ID NOS: 204966
;
; SEQ ID NO 92632
;
; LENGTH: 1671
;
; TYPE: DNA
;
; ORGANISM: Oryza sativa
;
; FEATURE:
;
; US-10-437-963-92632
;
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1
;
; US-10-437-963-92632

```

Query Match	21.9%	Score 301	DB 17	Length 1671
Best Local Similarity	67.5%	Pred. No. 2.6e-71		
Matches 585	Conservative 0	Mismatches 220	Indels 62	Gaps 9
Qy 113	AGGCCGCGCGCTTGCTCAGATGGGAAGTACATGCGCAAGTCACGGGGCGCGCAGGC	172		
Db				
1227	AGGGGCGCGACCGCGCGAGATGGGCAAGTACATGAGGAAGTTCACGGGGGGCCACGGGG	1168		
Qy 173	GCGGAGTGCCTCCGCGCTCGAGGTTAGCAGGTCGTGGCGCTCCGGACGAGGTCACAGTCC	232		
Db				
1167	GAGGAGTTGGCCGCCATGGAGGTCACGAGGTGGTTGGCGCTCCGCACAGGTCGAGTGC	1108		
Qy 233	GCGGCGCGACCGCGC-----GTGTCGCAAGGTCGCCCGCAGGAGG	274		
Db				
1107	GCAGCGCGCGGGCGCGACGACGAGGTCGAGGCGGCGCTCGGGCGTCCACCAG	1048		
Qy 275	AAGAGCGCGCGCGGGGAGCCTGCTGCGCGCTGAGCGCTGGTGGGACGGCGAAGC	334		
Db				
1047	AGGAGGAAGCGCTGCTGCCACGGCGGTCGTGGGACTACTCGCGCTGACGGCGGAGC	988		
Qy 335	TGTCATCATCACTGGTAGCCGATGCTGTTATGGCACGGCTCAGCGCAGCCGCTCG	394		
Db				
987	TGTCATCCTCAGCTGAGAGCCGATGCTGTTATGGCCCCCGC---GAGCGCGGGCCCG	931		
Qy 395	GTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGGCGCTGCAGGACACAGCGGGCGC	454		

Db 930 GCGCGAGGCTCCGTTTCTAGCGAGGCGCGGTTCCGGACACGAGCGCGCGCAT 871
QY 455 GCGCTCGCGCGCGGCTCTCGGTTCTCAGACACGCGGTCGTCGCTGAACTTGGGCTTG 514
Db 870 GCG--GCGGCTGGCTCTCGGTTCTCAGACACGCGCTGTC-----CGTGGACGCG 820
QY 515 GGGGGTCAGCGGGAGCACACCTCGCGCTCTAGACGCTCAGAGCTGCGGGAT 574
Db 819 GCGCTCAGACAGGAGCTCGGTCGCGTC---CGAGCTCGGAGGAGCAGCGAG 763
QY 575 CAGCTCTGCTGATGCTCTCGCGGAGCACTCGGGAGCGGCCAGACCCGAGAGG 634
Db 762 CATGTCGCGAGG-----GCTCCGAGGAGCTCGCGAGCGCGTGAACCGAGAGG 709
QY 635 CGAGACAGACGATCAGACCGCGGCGAGCGGAGCTCAGCGATCTGAGTCGATCTG 694
Db 708 AGAGAAACAACTCCATCAAGCTTTCTCCCGGAGAGGTCAGCGATCTGGAGTCGATCTG 649
QY 695 GCGGGSCACAAG-----CTGGCCGCTCGTACCGCGCGCAACGCGG 736
Db 648 GTGAGGACAGAGCGAGCGCTCCACTACCTTCTGCGGCAACAGCCTCAGCACAGAA 589
QY 737 GCTGCGAGCTGATCGTCCGCGCAGCACAGATCCAGAGTCTTCTCGCGCGCGCGGAG 796
Db 588 GCCAGCGCGCGAAGATTCCGCGCGCGCGGAGATCGAGCGGTTCTTCGCGCGCGCGAG 529
QY 797 GCGGCCAGGCAAGCGCTTCTTCCAAAGTACAACTTCACTTCTCGCGCGCGTGGCC 856
Db 528 GAGGCTGAGCGCAAGCGCTTCGCGCGCAAGTACAACTTCACTTCTCGCGCGTGGCC 469
QY 857 CTCGACGCGCGCGCGTTCAGTGGCGCGGTCGTCAGATCTGAGCTCTGAAGC--GAGCGTG 914
Db 468 CTCGAGC---CGTTCGTTTCAGTGGATCTCGGTCGTCAGCGCGGAGCTGAAGCGAG 412
QY 915 COTCGGTCGAAGGTGAAGTGAAG 941
Db 411 CGTGCAATTAAGCGAAGCTAGAAAG 385

RESULT 9

US-10-333-006-9
; Sequence 9, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz., Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-333-006-9

Query Match 21.8%; Score 299; DB 16; Length 1242;
Best Local Similarity 68.0%; Pred. No. 7,7e-71;
Matches 577; Conservative 0; Mismatches 210; Indels 62; Gaps 9;

QY 131 CAGATGGGAAGTACATGCGCAAGTCCAGGGGCGCCGAGGCGCGGAGGTCGCGCGGCTC 190
Db 15 CAGATGGGAAGTACATGAGGAAGTTTCAGGGGGGCCACCGGGGAGGAGTTGGCGCGCATG 74
QY 191 GAGGTTACCGAGGTCGTCGCGCTCCGAGACAGAGTCCAGGTCCGCGCGGCGAGCCGGCG-- 248
Db 75 GAGGTACCGAGGTTGTTGCGCTCCGAGACAGAGTCCAGGTCCGAGGTCCGCGCGGCGGCGG 134
QY 249 -----GTGTCGGAAGGTCGCGCCCGAGGAGGAAGGGCGCGCGGGGG 292
Db 135 AGCAGCACAAAGGTGAGGCGCGCTCGCGGGCTCCACAGAGGAGGAAGGGCTGCTG 194
QY 293 GAGCCTGCTCGCGCTGAGCGTGTGGGAGCGGCGGAGCTGTACATCCACTGCGT 352
Db 195 CCGACGCGCGTCTGTGGGACTACTCGCGCTGACGCGGAGGCTGCTACCTCCAGCTGAGG 254
QY 353 AGCGCATCTGTTTATGTCACCGCTCAGCGCGACCGCTCGGTTGACTCGTTCGGAACC 412
Db 255 AGCGCATCTGTTTATGTCGCGCGC---GAGGCGCGCGCGCGGAGGCTCGGTT 311
QY 413 CCGGTGAGGCTGTGATGGCGTGCAGACAGAGGCGCGCGCTCGCGCGCGGCTC 472
Db 312 GTAGCGAGGCGCGGTTTCCGGAACGAGCGCGCGCGCATGCG--GCGGCTGGCTC 368
QY 473 TCGCGTTGCTCCAGCACGCGCTCGTGTGAATTTGGGCTTGGGGGTTCAGCGCGGAGC 532
Db 369 TCGGTTGCTCCAGCACGCGCTCGT-----CGTGGACGCGCGCGCTCAGACAGGAGC 422
QY 533 CACACTGCGCTCTTACGACGCTGCAGAGGCTGGCGGGATCACGCTCTGTTGATGTC 592
Db 423 CTCGCTGCGCTC---CGAGCTCGCGAGGCGAGCGAGCATGTCTCCGAGG-- 474
QY 593 TCGCGCGAGGAACTTCGCGGAGGCGGCCAGACCGCGAGGCGAGAGCAGCGCATCG 652
Db 475 -GCTCGCGAGGAGCTTCGCGAGCGCGCTGACCGCGAGGAGAGAAACAATCCATCA 533
QY 653 AGCGGGGCGACGCGAGCTCAGCATCTGAGTCTGATCTCGCGGGGCGACAGA----- 707
Db 534 AGCTTCTCCCGCGAGGTGAGCATCTGGAGTCTGGATCTGGAGGACAGAGCGC 593
QY 708 -----CTGGCGCGCTCGCTTACCGCGGCAACCGCGCTCGGAGCTGTGCTG 754
Db 594 AGCGGTCACACTCTCTCGCGCAACAGCTCAGCACAGCAAGCCACGCGCGAAGATT 653
QY 755 CCGCAGCACAGAGATCCAGAGTTCTTCGCGCGCGCGCGAGCGCGCCAGGCGCAAGCGC 814
Db 654 CCGCGCGCGCGAGATCGAGGCTTCTTCGCGCGCGCGCGAGAGGCTGAGGCGCAAGCGC 713
QY 815 TTTGCTTCCAAGTACAACTTCGACTTCGTCGCGCGGCTCGCTCGACGCGCGCGCGCG 874
Db 714 TTCGCGCGCAAGTACAACTTCGAGCTGTTCGCGGCTGCCCTTCAGCGC---CGGTCGG 770
QY 875 TTCGAGTGGCGCGCGTGTGTGATCTGAAGC--GAGCGTCTGCTCCGCTGCAAGTGAA 932
Db 771 TTCGAGTGGACTCCGCTGTGTGATCTGAGCGAGCGCAAGCTGAAGCGCGCTGCAATTAAGCGAA 830
QY 933 GCTAGAAAG 941
Db 831 GCTAGAAAG 839

RESULT 10

US-10-767-701-24179
; Sequence 24179, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701

```
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24179
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Clone ID: 30165606
US-10-767-701-24179

Query Match
Best Local Similarity 19.9%; Score 272.4; DB 17; Length 510;
Matches 407; Conservative 0; Mismatches 71; Indels 42; Gaps 5;

Qy 12 GGACGCAAGCGGTGCGAGGAGAGCGCGCGAGGCGTGTGCGCTGTGGAGAGAA 71
Db 5 GAATCCAAAGCGCAAGCGGTGCGAGGAGAGCGCGCGAGGCGTGTGGAGAG 64
Qy 72 AAGAGAAAGAGAAACCGGCCCAAGACAAGCAAGCGAGAGCGCGCGGTTCGTC 131
Db 65 AGGAAAAGAGAAACCGGCCCAAGACAAGCAAGCGCGCGCGCGCGCGCGCG 118
Qy 132 AGATGGGAGTACATGC---GCAAGTGCAGGCGCGCGAGGCGCGAGGTTCGCGCG 188
Db 119 AGATGGGAGTACATGCAGCAAGTGCAGGCGCGCGCGCGCGCGCGCGCGAG 178
Qy 189 TCGAGGTTCAGCAGTGTGCGCGTCCGAGCAGAGTCCAGTCCGCGCGCGCGCGCG 248
Db 179 TCGAGGTTCAGCAGTGTGCGCGTCCGAGCAGAGTCCAGTCCGCGCGCGCGCG 238
Qy 249 GTGTCCGAGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 290
Db 239 GTGTCCGAGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 298
Qy 291 GGGAGCTGTCCGCGCGTGTGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCG 350
Db 299 CGGGGAGCGTCCGCGCGTGTGCGCGTGTGCGCGCGCGCGCGCGCGCGCG 358
Qy 351 GTAGCGCATCTGTTCATGCACCGCTCAGCGCGCGCGCGCGCGCGCGCGCG 410
Db 359 GAACCGCATCTGTTCATGCACCGCTCAGCGCGCGCGCGCGCGCGCGCGCG 409
Qy 411 CCGCGGTGAGGCTGTCTGATGCGCGTGCAGACAGCGCGCGCGCGCGCGCGCG 470
Db 410 CGCCCGGAGGCTGTCTGATGCGCGTGCAGACAGCGCGCGCGCGCGCGCGCG 463
Qy 471 TCTCGGTGTCTCAGCAGCGCGTGTCTGCTGTAAGTGGG 510
Db 464 TCTCGGTGTCTCAGCAGCGCGTGTCTGCTGTAAGTGGG 503

RESULT 11
US-10-333-006-5
; Sequence 5, Application US/10333006
; Publication No. US2004001926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US2004001926A111e S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n = A, T, C or G
US-10-333-006-5

Query Match
Best Local Similarity 16.0%; Score 219; DB 16; Length 568;
Matches 335; Conservative 0; Mismatches 36; Indels 25; Gaps 7;

Qy 563 GCTTGGCGGGATCAGCTCTGGTGGTGTCTCGCGCGGAGCACTCCGGAGCGGCCA 622
Db 2 GGTTCACGGGATCAGCTCTGGTGGTGTCTCGCGCGGAGCACTCCGGAGCGGCC 58
Qy 623 GACCGGAGAGGCGGAGAGAGAGCGCCCA---TCGAGCGCGCGCGCA---CGCGAGCTCAGC 676
Db 59 GACCGGAGAGGAGAGAGAGAGCGCCCA---TCGAGCGCGCGCGCA---CGCGAGCTCAGC 118
Qy 677 GATCTGAGTCTGATCTGCGCGG---GGCACAAGACTGCG---CCGTGCTACCGCG 727
Db 119 GATCTGAGTCTGATCTGCGCGG---GGCACAAGACTGCG---CCGTGCTACCGCG 178
Qy 728 GCAACGCGCGTCTGCGAGCTGATCGTCCGCGCAGCAGAGATCCAGGAGTCTTCGCG 787
Db 179 ACAACATCGGCTGCGAGCTGATCGTCCGCGCAGCAGAGATCCAGGAGTCTTCGCG 238
Qy 788 GCGCGCGAGGCGCGCGCGCGCGCGCGCTTTGCTTCCAAAGTACAACTTCGATTCGCG 847
Db 239 GCGCGCGAGGCGCGCGCGCGCGCGCGCTTTGCTTCCAAAGTACAACTTCGATTCGCG 298
Qy 848 GCGTGTGCGTCTGAGCGCGCGCGCGCGCGTTCGAGTGGCGCGCGTGGTTCAGCATCTGAGC 907
Db 299 GCGTGTGCGTCTGAGCGCGCGCGCGCGCGTTCGAGTGGCGCGCGTGGTTCAGCATCTGAGC 355
Qy 908 GAGCGTGTGCGTCTGAGCGCGCGCGCGCGCGTTCGAGTGGCGCGCGTGGTTCAGCATCTGAGC 943
Db 356 GAGCGTGTGCGTCTGAGCGCGCGCGCGCGCGTTCGAGTGGCGCGCGTGGTTCAGCATCTGAGC 387

RESULT 12
US-10-653-595-104
; Sequence 104, Application US/10653595
; Publication No. US20040048304A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027P1C1
; CURRENT APPLICATION NUMBER: US/10/653,595
; CURRENT FILING DATE: 2003-09-03
; PRIOR APPLICATION NUMBER: US 09/397945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080,312
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/078,578
```

;; PRIOR FILING DATE: 1998-03-19
;; Remaining Prior Application data removed - See File Wrapper or PALM.

;; NUMBER OF SEQ ID NOS: 470
;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 104
;; LENGTH: 2015

;; TYPE: DNA
;; ORGANISM: Homo sapiens

;; FEATURE:
;; NAME/KEY: SITE

;; LOCATION: (3)

;; OTHER INFORMATION: n equals a,t,g, or c

;; FEATURE:
;; NAME/KEY: SITE

;; LOCATION: (9)

;; OTHER INFORMATION: n equals a,t,g, or c

;; FEATURE:
;; NAME/KEY: SITE

;; LOCATION: (1981)

;; OTHER INFORMATION: n equals a,t,g, or c

;; FEATURE:
;; NAME/KEY: SITE

;; LOCATION: (1990)

;; OTHER INFORMATION: n equals a,t,g, or c

;; FEATURE:
;; NAME/KEY: SITE

;; LOCATION: (2002)

;; OTHER INFORMATION: n equals a,t,g, or c

US-10-653-595-104

Query Match 8.4%; Score 114.8; DB 13; Length 2015;
Best Local Similarity 86.8%; Pred. No. 1.8e-20;

Matches 125; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 1228 TGSITGGCTTGCAGGAAGATCTGGTTGTCTCAAAAAAAAAAAAAAAAAAAAAA 1287

Db 1514 TGTGTAGGATAAAATAAACCATCTTGTATGGGAAAAAAAAAAAAAAAAAAAA 1573

QY 1288 AAGGGCGCGCTCTAGAGATCCCAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCT 1347

Db 1574 AAGGGCGCGCTCTAGAGATCCCAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCT 1633

QY 1348 TCTATAGTGTACCTAAATTCATT 1371

Db 1634 TCTATAGTGTACCTAAATTCATT 1657

RESULT 13

US-09-397-945-104

;; Sequence 104, Application US/09397945

;; Publication No. US20030065139A1

;; GENERAL INFORMATION:

;; APPLICANT: Human Genome Sciences, Inc. et al.

;; TITLE OF INVENTION: 95 Human secreted proteins

;; FILE REFERENCE: P2027P1

;; CURRENT APPLICATION NUMBER: US/09/397,945

;; PRIOR FILING DATE: 1999-09-17

;; PRIOR APPLICATION NUMBER: PCT/US99/05804

;; PRIOR FILING DATE: 1999-03-18

;; PRIOR APPLICATION NUMBER: 60/078,566

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/078,576

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/078,573

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/078,574

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/078,579

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/080,314

;; PRIOR FILING DATE: 1998-04-01

;; PRIOR APPLICATION NUMBER: 60/080,312

;; PRIOR FILING DATE: 1998-04-01

;; PRIOR APPLICATION NUMBER: 60/078,578

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/078,581

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/078,577

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/078,563

;; PRIOR FILING DATE: 1998-03-19

;; PRIOR APPLICATION NUMBER: 60/080,313

;; PRIOR FILING DATE: 1998-04-01

;; NUMBER OF SEQ ID NOS: 470

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 104

;; LENGTH: 2015

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:
;; NAME/KEY: SITE

;; LOCATION: (3)

;; OTHER INFORMATION: n equals a,t,g, or c

;; NAME/KEY: SITE

;; LOCATION: (9)

;; OTHER INFORMATION: n equals a,t,g, or c

;; NAME/KEY: SITE

;; LOCATION: (1981)

;; OTHER INFORMATION: n equals a,t,g, or c

;; NAME/KEY: SITE

;; LOCATION: (1990)

;; OTHER INFORMATION: n equals a,t,g, or c

;; NAME/KEY: SITE

;; LOCATION: (2001)

;; OTHER INFORMATION: n equals a,t,g, or c

;; NAME/KEY: SITE

;; LOCATION: (2002)

;; OTHER INFORMATION: n equals a,t,g, or c

US-09-397-945-104

Query Match 8.4%; Score 114.8; DB 13; Length 2015;
Best Local Similarity 86.8%; Pred. No. 1.8e-20;

Matches 125; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 1228 TGSITGGCTTGCAGGAAGATCTGGTTGTCTCAAAAAAAAAAAAAAAAAAAAA 1287

Db 1514 TGTGTAGGATAAAATAAACCATCTTGTATGGGAAAAAAAAAAAAAAAAAAAA 1573

QY 1288 AAGGGCGCGCTCTAGAGATCCCAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCT 1347

Db 1574 AAGGGCGCGCTCTAGAGATCCCAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCT 1633

QY 1348 TCTATAGTGTACCTAAATTCATT 1371

Db 1634 TCTATAGTGTACCTAAATTCATT 1657

RESULT 14

US-09-764-846-83

;; Sequence 83, Application US/09764846

;; Patent No. US20020102638A1

;; GENERAL INFORMATION:

;; APPLICANT: Rosen et al.

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

;; FILE REFERENCE: PT212

;; CURRENT APPLICATION NUMBER: US/09/764,846

;; CURRENT FILING DATE: 2001-01-17

;; Prior application data removed - consult PALM or file wrapper

;; NUMBER OF SEQ ID NOS: 348

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 83

;; LENGTH: 3313

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3264)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3308)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3309)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-846-83

Query Match 8.1%; Score 110.8; DB 9; Length 3313;
Best Local Similarity 89.4%; Pred. No. 2.9e-19;
Matches 118; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1240 AAGAAAGATCTGGTTTGTCTCAAAAAAAAAAAAAAAAAAGGGCGCGC 1299
Db 3107 AATAAAACATTTTGGATTTTCAAAAAAAAAAAAAAAAAAGGGCGCGC 3166
Qy 1300 TCTAGAGATCCAAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCTTCTATAGTGTC 1359
Db 3167 TCTANAGATCCAAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCTTCTATAGTGTC 3226
Qy 1360 CCTAAATTCATT 1371
Db 3227 CCTAAATTCATT 3238

RESULT 15
US-10-091-483-83
; Sequence 83, Application US/10091483
; Publication No. US20030049650A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ12C1
; CURRENT APPLICATION NUMBER: US/10/091,483
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 348
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 3313
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3264)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3308)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3309)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-483-83

Query Match 8.1%; Score 110.8; DB 15; Length 3313;

Best Local Similarity 89.4%; Pred. No. 2.9e-19;
Matches 118; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1240 AAGAAAGATCTGGTTTGTCTCAAAAAAAAAAAAAAAAAAGGGCGCGC 1299
Db 3107 AATAAAACATTTTGGATTTTCAAAAAAAAAAAAAAAAAAGGGCGCGC 3166
Qy 1300 TCTAGAGATCCAAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCTTCTATAGTGTC 1359
Db 3167 TCTANAGATCCAAAGCTTACGTACGCGTGCATGCGAGTCATAGCTCTTCTATAGTGTC 3226
Qy 1360 CCTAAATTCATT 1371
Db 3227 CCTAAATTCATT 3238

Search completed: October 2, 2004, 06:02:18
Job time : 713 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 05:36:42 ; Search time 3331 Seconds
(without alignments)
12299.883 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtcggagcaag.....agtgcacctaattcattc 1372

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmd:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estcom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389	28.4	526	14	CD444277 EL01N0438
2	197	14.4	890	29	CG218373 OG1DJ10TV
3	197	14.4	1010	28	CC390569 PUHBT76TB
4	193	14.1	753	29	CC729673 OGUBY35TH

5	193	14.1	791	29	CG324885	CG324885	OGXDN04TV
6	193	14.1	825	29	CG265472	CG265472	OG2CM60TH
7	193	14.1	917	29	CG265481	CG265481	OG2CM60TV
8	178	13.0	547	28	BH777311	BH777311	Fzmb013f0
9	167	12.2	884	29	CG223284	CG223284	OGWAU17TH
10	167	12.2	885	29	CG324875	CG324875	OGXDN04TH
11	167	12.2	887	29	CG223296	CG223296	OGWAU17TV
12	160	11.7	707	11	AY110319	AY110319	Zea mays
13	152	11.1	986	28	CC002416	CC002416	FUGKB53TB
14	152	11.1	987	28	CC002419	CC002419	FUGKB53TD
15	143	10.4	645	29	CG136784	CG136784	PUIDC25TB
16	131	9.5	225	14	CD957606	CD957606	SCK 144 G
17	119	8.7	505	14	CA402478	CA402478	EL01N0438
18	112	8.2	741	29	CC683397	CC683397	OGUHL76TV
19	112	8.2	790	28	CC390573	CC390573	PUHBT76TD
20	112	8.2	835	29	CG218361	CG218361	OG1DJ10TH
21	109	7.9	669	12	BG924526	BG924526	HNC35-1-D
22	108	7.9	172	14	CF643461	CF643461	D90 H03 F
23	108	7.9	397	12	BG898272	BG898272	HOA3-1-C
24	108	7.9	402	12	BG897464	BG897464	HOA14-1-D
25	108	7.9	404	12	BG896735	BG896735	HOA49-1-F
26	108	7.9	433	12	BG898370	BG898370	HOA8-1-E4
27	108	7.9	441	12	BG927036	BG927036	HNC35-1-E
28	108	7.9	473	12	BG898628	BG898628	HOA3-1-B
29	108	7.9	480	12	BG928051	BG928051	HNC46-1-H
30	108	7.9	559	12	BG924503	BG924503	HNC41-1-F
31	108	7.9	583	12	BG927127	BG927127	HNC17-1-G
32	108	7.9	586	12	BG926479	BG926479	HNC62-1-B
33	108	7.9	591	12	BG928233	BG928233	HNC47-1-F
34	108	7.9	595	12	BG924343	BG924343	HNC26-1-F
35	108	7.9	608	12	BG896941	BG896941	HOA59-1-B
36	108	7.9	609	12	BG924870	BG924870	HNC89-1-F
37	108	7.9	609	12	BG929210	BG929210	HNC59-1-G
38	108	7.9	618	12	BG928558	BG928558	HNC68-1-E
39	108	7.9	620	12	BG925002	BG925002	HNC39-1-B
40	108	7.9	628	12	BG928414	BG928414	HNC65-1-H
41	108	7.9	631	12	BG927739	BG927739	HNC44-1-G
42	108	7.9	632	12	BG926073	BG926073	HNC23-1-F
43	108	7.9	633	12	BG899001	BG899001	HOA58-1-B
44	108	7.9	634	12	BG899738	BG899738	HOA39-1-E
45	108	7.9	635	12	BG924853	BG924853	HNC36-1-A

ALIGNMENTS

RESULT 1
CD444277
LOCUS CD444277 526 bp mRNA linear EST 03-JUN-2003
DEFINITION EL01N0438A02.b Endosperm_4 Zea mays cDNA, mRNA sequence.
ACCESSION CD444277
VERSION CD444277.1 GI:31359920
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 526)
AUTHORS Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
Messing, J.
TITLE Sequencing of the maize endosperm ESTs
JOURNAL Unpublished (2002)
COMMENT Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
FEATURES Location/Qualifiers
source 1..526


```

Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 AGCGGAGAGACGACGCTCGAGCCGGGCGCAGCGGAGCTAGCGATCTGGAGTCGGAT 691
Db 177 AGCGGAGAGACGACGCTCGAGCCGGGCGCAGCGGAGCTAGCGATCTGGAGTCGGAT 236
QY 692 CTGGCGGGGCACAAAGACTGGCCGCTCGCTACCGCGGCAACCGCGCTGGAGCTGATC 751
Db 237 CTGGCGGGGCACAAAGACTGGCCGCTCGCTACCGCGGCAACCGCGCTGGAGCTGATC 296
QY 752 GTGCCGCCAGCACACAGATCAGGAGTTCTTCGCCGCCGCCGCCGCCGCCGCCGCCCAAG 811
Db 297 GTGCCGCCAGCACACAGATCAGGAGTTCTTCGCCGCCGCCGCCGCCGCCGCCGCCCAAG 356
QY 812 CGCTTTGCTTCCAAGTA 828
Db 357 CGCTTTGCTTCCAAGTA 373

RESULT 4
CC729673/c
LOCUS
DEFINITION
  OGBY35TH ZM 0.7 1.5 KB Zea mays genomic clone ZMBMa0404F21,
  genomic survey sequence.
ACCESSION
  CC729673
VERSION
  CC729673.1 GI:32148606
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 753)
REFERENCE
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OGBY35TV
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
FEATURES
  source
    1..753
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMBMa0404F21"
    /clone_lib="ZM 0.7 1.5 KB"
    /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"
ORIGIN
  Query Match 14.1%; Score 193; DB 29; Length 753;
  Best Local Similarity 100.0%; Pred. No. 9.3e-29;
  Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 GAGAGACGACGCTCGAGCCGGGCGCAGCGGAGCTCAGCGATCTGGAGTCGGATCTGG 695
Db 406 GAGAGACGACGCTCGAGCCGGGCGCAGCGGAGCTCAGCGATCTGGAGTCGGATCTGG 347
QY 696 CGGGGCACAGACTGGCCGCTCGCTACCGCGGCAACCGCGCTGGAGCTGATCGTGC 755
Db 346 CGGGGCACAGACTGGCCGCTCGCTACCGCGGCAACCGCGCTGGAGCTGATCGTGC 287
QY 756 CGCCAGCACACGAGATCCAGAGTTCTTCGCCGCCGCCGCCGCCGCCGCCGCCCAAGCGCT 815
  1 (bases 1 to 753)
  825 bp DNA linear GSS 25-AUG-2003
  OGBY35TH ZM 0.7 1.5 KB Zea mays genomic clone ZMBMa0764I23,
  genomic survey sequence.

```

```

Db 286 CGCCAGCACACGAGATCCAGGAGTTCTTCGCCGCCGCCGCCGCCGCCGCCCAAGCGCT 227
QY 816 TTGCTTCCAAGTA 828
Db 226 TTGCTTCCAAGTA 214

RESULT 5
CG324885
LOCUS
DEFINITION
  CGXND04TV ZM 0.7 1.5 KB Zea mays genomic clone ZMBMa0661B08,
  genomic survey sequence.
ACCESSION
  CG324885
VERSION
  CG324885.1 GI:34242151
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 791)
REFERENCE
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OXND04TH
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
FEATURES
  source
    1..791
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMBMa0661B08"
    /clone_lib="ZM 0.7 1.5 KB"
    /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"
ORIGIN
  Query Match 14.1%; Score 193; DB 29; Length 791;
  Best Local Similarity 100.0%; Pred. No. 8.9e-29;
  Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 GAGAGACGACGCTCGAGCCGGGCGCAGCGGAGCTCAGCGATCTGGAGTCGGATCTGG 695
Db 200 GAGAGACGACGCTCGAGCCGGGCGCAGCGGAGCTCAGCGATCTGGAGTCGGATCTGG 259
QY 696 CGGGGCACAGACTGGCCGCTCGCTACCGCGGCAACCGCGCTGGAGCTGATCGTGC 755
Db 260 CGGGGCACAGACTGGCCGCTCGCTACCGCGGCAACCGCGCTGGAGCTGATCGTGC 319
QY 756 CGCCAGCACACGAGATCCAGGAGTTCTTCGCCGCCGCCGCCGCCGCCGCCCAAGCGCT 815
  1 (bases 1 to 791)
  825 bp DNA linear GSS 25-AUG-2003
  OGBY35TH ZM 0.7 1.5 KB Zea mays genomic clone ZMBMa0764I23,
  genomic survey sequence.

```

ACCESSION CG265472
 VERSION CG265472.1 GI:34177613
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 825)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OG2CM60TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1..825
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0764123"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
 ORIGIN
 Query Match 14.1%; Score 193; DB 29; Length 825;
 Best Local Similarity 100.0%; Pred. No. 8.5e-29;
 Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 636 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
 Db 816 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 757
 QY 696 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 755
 Db 756 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 697
 QY 756 CGCAGACACAGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCT 815
 Db 696 CGCCAGCACACGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGCT 637
 QY 816 TTGCTTCCAGTA 828
 Db 636 TTGCTTCCAGTA 624
 RESULT 7
 CG265481
 LOCUS CG265481 917 bp DNA linear GSS 25-AUG-2003
 DEFINITION OG2CM60TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0764123, genomic survey sequence.
 ACCESSION CG265481
 VERSION CG265481.1 GI:34177622
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 917)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OG2CM60TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1..825
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0764123"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
 ORIGIN
 Query Match 14.1%; Score 193; DB 29; Length 825;
 Best Local Similarity 100.0%; Pred. No. 8.5e-29;
 Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 636 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
 Db 816 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 757
 QY 696 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 755
 Db 756 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 697
 QY 756 CGCAGACACAGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCT 815
 Db 696 CGCCAGCACACGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGCT 637
 QY 816 TTGCTTCCAGTA 828
 Db 636 TTGCTTCCAGTA 624
 RESULT 7
 CG265481
 LOCUS CG265481 917 bp DNA linear GSS 25-AUG-2003
 DEFINITION OG2CM60TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0764123, genomic survey sequence.
 ACCESSION CG265481
 VERSION CG265481.1 GI:34177622
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 917)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OG2CM60TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1..825
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0764123"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
 ORIGIN
 Query Match 14.1%; Score 193; DB 29; Length 917;
 Best Local Similarity 100.0%; Pred. No. 7.7e-29;
 Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 636 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
 Db 266 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 325
 QY 696 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 755
 Db 326 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 385
 QY 756 CGCAGACACAGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCT 815
 Db 386 CGCCAGCACACGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGCT 445
 QY 816 TTGCTTCCAGTA 828
 Db 446 TTGCTTCCAGTA 458
 RESULT 8
 BH777311/c
 LOCUS BH777311 547 bp DNA linear GSS 28-MAR-2002
 DEFINITION fzm013f023f06f0 fzm filtered library Zea mays genomic clone
 fzm013f023f06 5', genomic survey sequence.
 ACCESSION BH777311
 VERSION BH777311.1 GI:19779747
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 547)
 Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
 Genefresh, methylation filtered genomic sequences from maize
 Unpublished (2002)
 Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: fzm013f023 row: f column: 06
 Seq primer: M13 forward
 Class: shotgun
 High quality sequence stop: 547.
 Location/Qualifiers
 1..547

TITLE Consortium for Maize Genomics
 JOURNAL Unpublished (2002)
 COMMENT Other GSSs: OG2CM60TH
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1..917
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0764123"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
 ORIGIN
 Query Match 14.1%; Score 193; DB 29; Length 917;
 Best Local Similarity 100.0%; Pred. No. 7.7e-29;
 Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 636 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
 Db 266 GAGAGACACGCCATCGAGCGCGGCGGACGCGAGCTCAGCGATCTGGAGTCGGATCTGG 325
 QY 696 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 755
 Db 326 CGGGGCAACAGACTGGCCCGTCGCTACCGCGGCGCAACCGCGGCTCGGAGCTGATCGTGC 385
 QY 756 CGCAGACACAGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCT 815
 Db 386 CGCCAGCACACGAGATCCAGAGTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGCT 445
 QY 816 TTGCTTCCAGTA 828
 Db 446 TTGCTTCCAGTA 458
 RESULT 8
 BH777311/c
 LOCUS BH777311 547 bp DNA linear GSS 28-MAR-2002
 DEFINITION fzm013f023f06f0 fzm filtered library Zea mays genomic clone
 fzm013f023f06 5', genomic survey sequence.
 ACCESSION BH777311
 VERSION BH777311.1 GI:19779747
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 547)
 Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
 Genefresh, methylation filtered genomic sequences from maize
 Unpublished (2002)
 Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: fzm013f023 row: f column: 06
 Seq primer: M13 forward
 Class: shotgun
 High quality sequence stop: 547.
 Location/Qualifiers
 1..547

```

/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="Mol17"
/db_xref="taxon:4577"
/clone="fzmb013f023f06"
/clone_lib="fzmb filtered library"
/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to
5 kb fraction, ligated into HincII-digested pBCSK(-)
vector and electroporated into E. coli cells."

ORIGIN
Query Match 13.0%; Score 178; DB 28; Length 547;
Best Local Similarity 100.0%; Pred. No. 1e-25;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 AGCGAGAGACGACGCGGCGCAGCGGAGCTCAGCGATCGGATCGGAT 691
Db 294 AGCGAGAGACGACGCGGCGCAGCGGAGCTCAGCGATCGGATCGGAT 235
QY 692 CTGGCGGGCACAGACTGCGCTACCGCGGCGCAACGCGGCTGGGAGCTGATC 751
Db 234 CTGGCGGGCACAGACTGCGCTACCGCGGCGCAACGCGGCTGGGAGCTGATC 175
QY 752 GTCCGCCGACACAGATCCAGGAGTCTTCGCGCGCGCGGCGGCGGCGGCGCA 809
Db 174 GTCCGCCGACACAGATCCAGGAGTCTTCGCGCGCGCGGCGGCGGCGGCGCA 117

CG223284 884 bp DNA linear GSS 22-AUG-2003
OGWAU177M 0.7_1.5_KB Zea mays genomic clone ZMMBMA0524C09,
genomic survey sequence.
CG223284
CG223284.1 GI:34123172
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 884)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGWAU17TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..884
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0524C09"
/clone_lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 12.2%; Score 167; DB 29; Length 884;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 999 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCCACCAATTCATCTCGTCCAATGC 1058
Db 867 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCCACCAATTCATCTCGTCCAATGC 808
QY 1059 ATGCTTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTTACTTACTT 1118
Db 807 ATGCTTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTTACTTACTT 748
QY 1119 ATTACAGAGGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 1165
Db 747 ATTACAGAGGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 701

RESULT 10
CG324875/c
LOCUS
DEFINITION
CG324875 885 bp DNA linear GSS 26-AUG-2003
genomic survey sequence.
CG324875
CG324875.1 GI:34242141
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 885)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGXDN04TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..885
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0661B08"
/clone_lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 12.2%; Score 167; DB 29; Length 885;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 999 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCCACCAATTCATCTCGTCCAATGC 1058
Db 414 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCCACCAATTCATCTCGTCCAATGC 355
QY 1059 ATGCTTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTTACTTACTT 1118
Db 354 ATGCTTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTTACTTACTT 295
QY 1119 ATTACAGAGGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 1165
Db 294 ATTACAGAGGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 248

RESULT 11
CG223284

```

```

LOCUS      CG223296                      887 bp      DNA      linear      GSS 22-AUG-2003
DEFINITION  OGMWU17TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0524C09,
GENOMIC SURVEY SEQUENCE.
ACCESSION   CG223296
VERSION     CG223296.1      GI:34123184
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 887)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other GSSs: OGMWU17TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES    Location/Qualifiers
            source          1..887
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /strain="B73"
                        /db_xref="taxon:4577"
                        /clone="ZMMBMA0524C09"
                        /clone_lib="ZM 0.7 1.5 KB"
                        /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                        methylation filtered genomic DNA library"
ORIGIN
Query Match      12.2%; Score 167; DB 29; Length 887;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 999 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAAACCACCACTTCATCTCGTCCAAATGC 1058
DB 296 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAAACCACCACTTCATCTCGTCCAAATGC 355
QY 1059 ATGCTTCGCTTTCTCTGAGCTACAGAGCGTAGTTATTATTAGTACTACTTTACTTT 1118
DB 356 ATGCTTCGCTTTCTCTGAGCTACAGAGCGTAGTTATTATTAGTACTACTTTACTTT 415
QY 1119 ATTACAGAGTATCTTGACCCCGATAGATCAATCGCTTACTGTGTA 1165
DB 416 ATTACAGAGTATCTTGACCCCGATAGATCAATCGCTTACTGTGTA 462

RESULT 12
LOCUS      AY110319                      707 bp      mRNA      linear      HTC 17-OCT-2002
DEFINITION  Zea mays CL528-1 mRNA sequence.
ACCESSION   AY110319
VERSION     AY110319.1      GI:21214635
KEYWORDS    HTC.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 707)
AUTHORS     Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
            Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
            Unpublished (2002)

```

```

REFERENCE   2 (bases 1 to 707)
AUTHORS     Coe,E.H.
            Direct Submission
            Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT      If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSI, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Scimable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES    Location/Qualifiers
            source          1..707
                        /organism="Zea mays"
                        /mol_type="mRNA"
                        /db_xref="maizedb:632312"
                        /db_xref="taxon:4577"
                        /clone_lib="Maize Mapping Project/DuPont Consensus
                        Library"
                        /note="this sequence is part of a project of EST
                        assemblies resulting from the application of public
                        contigs to seed DuPont contigs; this resource was
                        assembled by DuPont as part of a collaboration for the
                        overgo addressing of BACs in conjunction with the Maize
                        Mapping Project"
ORIGIN
Query Match      11.7%; Score 160; DB 11; Length 707;
Best Local Similarity 100.0%; Pred. No. 2.4e-22;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1006 AAGCAGCTTTATATAGCCTAAGCTAAACCACCACTTCATCTCGTCCAAATGCCTT 1065
DB 476 AAGCAGCTTTATATAGCCTAAGCTAAACCACCACTTCATCTCGTCCAAATGCCTT 535
QY 1066 GCTTTTCTCGAGCTACAGAGCGTAGTTATTATTAGTACTACTTTACTTATTCAGA 1125
DB 536 GCTTTTCTCGAGCTACAGAGCGTAGTTATTATTAGTACTACTTTACTTATTCAGA 595
QY 1126 GGTATCTTGACCCCGATAGATCAATCGCTTACTGTGTA 1165
DB 596 GGTATCTTGACCCCGATAGATCAATCGCTTACTGTGTA 635

RESULT 13
LOCUS      CC002416                      986 bp      DNA      linear      GSS 31-MAR-2003
DEFINITION  PUGKB53TB ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA402110,
            genomic survey sequence.
ACCESSION   CC002416
VERSION     CC002416.1      GI:29380976
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 986)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Other GSSs: PUGKB53TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 04:19:11 ; Search time 562 Seconds
(without alignments)
10371.046 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccagcgccgagcagcaag.....agtgtaacctaattcattc 1372

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 3373863 seqs, 2124099041 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_29Jan04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1372	100.0	1372	7	ABV74603
2	322	23.5	723	3	AAA95281
3	108	7.9	149	4	AA102569
4	108	7.9	1316	3	AAC59966
5	108	7.9	1557	3	AAC99034
6	108	7.9	1557	5	AAS00830
7	108	7.9	2116	3	AAC99033
8	108	7.9	2116	5	AAS00847
9	106	7.7	701	4	AAH33453
10	105	7.7	1928	6	ABK69104
11	104	7.6	2152	4	AAD05156
12	103	7.5	845	6	ABQ54253
13	103	7.5	1088	4	ABD13354
14	99	7.2	2049	6	ABL30818
15	98	7.1	570	7	ABZ55039
16	98	7.1	2657	3	AAC69567
17	98	7.1	2657	7	ADA39657
18	98	7.1	2657	9	ADC73401
19	97	7.1	544	4	AAH35303
20	97	7.1	652	6	ABQ55965
21	97	7.1	2454	3	AAC69584
22	97	7.1	2454	7	ADA40257
23	97	7.1	2454	9	ADC73781

24	96	7.0	377	4	AAI98573	Xai98573 Human exc
25	96	7.0	377	4	AAI64026	Xai64026 Human bla
26	96	7.0	452	6	ABQ55723	Abq55723 Human ova
27	96	7.0	762	6	ABQ55923	Abq55923 Human ova
28	96	7.0	894	9	ADB47734	Adb47734 Novel hum
29	96	7.0	973	8	AAD56374	Ad56374 Human sec
30	96	7.0	1138	3	AAFI6267	Xafi6267 Human pro
31	96	7.0	1631	4	AAK64510	Xak64510 Human imm
32	96	7.0	1958	4	AAH35000	Aah35000 Human col
33	96	7.0	1958	6	ABL90702	Ab190702 Human pol
34	96	7.0	1958	6	ABQ54929	Abq54929 Human ova
35	96	7.0	3240	4	AAF75345	Aaf75345 Human TGF
36	95	6.9	395	4	AAH35297	Aah35297 Human col
37	95	6.9	498	3	AAC98321	Aac98321 Human col
38	95	6.9	606	6	ABQ55926	Abq55926 Human ova
39	95	6.9	611	4	AAK88301	Aak88301 Human dig
40	95	6.9	611	5	AAS39445	Aas39445 cDNA enco
41	95	6.9	611	8	ADB32171	Adb32171 Human nov
42	95	6.9	657	6	ABQ55900	Abq55900 Human ova
43	95	6.9	661	6	ABQ55924	Abq55924 Human ova
44	95	6.9	670	6	ABQ54585	Abq54585 Human ova
45	95	6.9	684	3	AAC98310	Aac98310 Human col

ALIGNMENTS

RESULT 1
ABV74603
ID ABV74603 standard; DNA; 1372 BP.
XX
AC ABV74603;
XX
DT 21-FEB-2003 (first entry)
XX
DE Maize CKI_B coding sequence.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication; gene; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 134..904
FT /*tag= a
FT /product= "CKI_B"
XX
PN WO200281623-A2.
XX
PD 17-OCT-2002.
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
WPI; 2003-058511/05.
P-PSDB; ABB98757.
XX
PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX
PS Claim 1; Page 64-65; 69pp; English.
CC The present sequence is the coding sequence (I) for maize cyclin-
dependent kinase inhibitor (CKI), CKI_B. (I) is useful for modulating the

activity of cyclin-dependent kinase (CDK) in a plant such as a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, oil-seed Brassica and millet. Modulating the activity of CDK, preferably modulating downward is useful for providing differential growth in a plant, especially a positive growth advantage and modulating CDK activity upward is useful for increasing crop yield, root size, plant growth, tassel size and/or ear size. Modulating CDK activity is also useful for conferring male sterility and for improving transformation frequencies by increasing the number of cells in cell division. CDK activity can also be modulated for modulating endoreduplication in the endosperm of corn, sorghum, wheat, rice, barley, and millet, where the promoter used is an endosperm-preferred promoter. The cell numbers are modulated in one or more tissues of a plant, comprising root, seed, tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves, stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm, pericarp, meristematic or leaf cells. (1) is also useful for identifying maize CKI interacting proteins, by aducting the nucleic acid sequence to a second nucleic acid sequence encoding a DNA-binding domain

XX SQ Sequence 1372 BP; 292 A; 396 C; 442 G; 241 T; 0 U; 1 Other;

Query Match 100.0%; Score 1372; DB 7; Length 1372;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCACCGCTCCGGACGCAAGCGGTGACGAGCGCGCGCAGAGCGGTGTGGCCCTGT 60
 Db 1 CCCACCGCTCCGGACGCAAGCGGTGACGAGCGCGCGCAGAGCGGTGTGGCCCTGT 60

61 GGGAGAGAAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAG 120
 Db 61 GGGAGAGAAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAG 120

121 GGGCTGCGTCAGATGGGAGAGTACATCGCAAGTGCAGAGCGCGCGCAGAGCGGAGGT 180
 Db 121 GGGCTGCGTCAGATGGGAGAGTACATCGCAAGTGCAGAGCGCGCGCAGAGCGGAGGT 180

181 CGCCGCGCTCAGAGTTACGCAAGTGTCCGCGGTCCGAGCGAGGTCCAGGTCCGCGCGGC 240
 Db 181 CGCCGCGCTCAGAGTTACGCAAGTGTCCGCGGTCCGAGCGAGGTCCAGGTCCGCGCGGC 240

241 GACCGCGGTGTCCGAAAGTGTCCGCGGTCCGCGGTCCGCGGTCCGCGGTCCGCGGTG 300
 Db 241 GACCGCGGTGTCCGAAAGTGTCCGCGGTCCGCGGTCCGCGGTCCGCGGTCCGCGGTG 300

301 TGCCGCGTGTGAGCGGTGTGGGAGCGCGGAGGTGTACATCCACCTGCGTAGCCGAT 360
 Db 301 TGCCGCGTGTGAGCGGTGTGGGAGCGCGGAGGTGTACATCCACCTGCGTAGCCGAT 360

361 GCTGTTATGTCACCGCTCAGCGCGCAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 361 GCTGTTATGTCACCGCTCAGCGCGCAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

421 GGTGCTGATGCGGTGACGAGCAGAGCGCGCGCGGTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 421 GGTGCTGATGCGGTGACGAGCAGAGCGCGCGCGGTGCTGCTGCTGCTGCTGCTGCTG 480

481 CTCAGACACGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 481 CTCAGACACGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

541 CCGTCTCTACGACGCTGACGAGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCG 600
 Db 541 CCGTCTCTACGACGCTGACGAGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCG 600

601 GAGCAACTCTCCGGAGCG 660
 Db 601 GAGCAACTCTCCGGAGCG 660

661 GCACGGCGAGTCTCAGCGATCTGGAGTCTGGAGTCTGGAGTCTGGAGTCTGGAGTCT 720
 Db 661 GCACGGCGAGTCTCAGCGATCTGGAGTCTGGAGTCTGGAGTCTGGAGTCTGGAGTCT 720

721 ACCGCGCGCAACGCGCGGTGCGAGCTGATCGTCCGCCACGACACGAGATCCAGAGTT 780
 Db 721 ACCGCGCGCAACGCGCGGTGCGAGCTGATCGTCCGCCACGACACGAGATCCAGAGTT 780

781 CTTGCGCGCCCGCGAGCGCGCCGAGCGCTTGGCTTCCAGAGTACACTTCGACTT 840
 Db 781 CTTGCGCGCCCGCGAGCGCGCCGAGCGCTTGGCTTCCAGAGTACACTTCGACTT 840

841 CGTCGCGCGCTGCCCTCCGACGCGCGCGCGGTTCGAGTGGCGCGCGGTTCGAGCAT 900
 Db 841 CGTCGCGCGCTGCCCTCCGACGCGCGCGCGGTTCGAGTGGCGCGCGGTTCGAGCAT 900

901 CTGAAGCGAGCGTGTGCTGCAAGGTGAAGCTAGAAAGAGAAAGAGTCCCGCCGCC 960
 Db 901 CTGAAGCGAGCGTGTGCTGCAAGGTGAAGCTAGAAAGAGAAAGAGTCCCGCCGCC 960

961 CCCCCCGCCCAACAAACATACGAGAGAGAGAAACCAAAACAATTAAGCAGCTTATATA 1020
 Db 961 CCCCCCGCCCAACAAACATACGAGAGAGAGAAACCAAAACAATTAAGCAGCTTATATA 1020

1021 GCCTAAGCTAACCAACCACTTCCTGTCCTCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 Db 1021 GCCTAAGCTAACCAACCACTTCCTGTCCTCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1080

1081 TAGCAGGAGCGTGTATTTAGTACTACTTACTTACTTACTTACTTACTTACTTACTTACT 1140
 Db 1081 TAGCAGGAGCGTGTATTTAGTACTACTTACTTACTTACTTACTTACTTACTTACTTACT 1140

1141 GATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGATCTCTTAGTGTGAGTTTATC 1200
 Db 1141 GATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGATCTCTTAGTGTGAGTTTATC 1200

1201 GTCTTAATTTATTTACTGTACAGAGCTTGTGCTTCAAGAAAGATCTGTTTGTCT 1260
 Db 1201 GTCTTAATTTATTTACTGTACAGAGCTTGTGCTTCAAGAAAGATCTGTTTGTCT 1260

1261 CAAAAAAG 1320
 Db 1261 CAAAAAAG 1320

1321 TACGCGTGCATGCGAGCTGATAGCTCTTCTATAGTGTACCTTAATTCATTC 1372
 Db 1321 TACGCGTGCATGCGAGCTGATAGCTCTTCTATAGTGTACCTTAATTCATTC 1372

RESULT 2
 AAA95281 standard; cDNA; 723 BP.
 ID AAA95281
 AC AAA95281;
 DT 17-JAN-2001 (first entry)
 DE Corn cyclin-dependent kinase inhibitor coding sequence #3.
 KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 cell growth; herbicide; ss.
 OS Zea mays.
 FH Key
 CDS Location/Qualifiers
 3..380
 /*tag= a
 /product= "CDKI"
 /partial
 XX WO200060087-A2.
 PN 12-OCT-2000.
 PD 06-APR-2000; 2000WO-US009106.
 XX 07-APR-1999; 99US-0128192P.
 PR


```
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
XX P-PSDB; AAB26250.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
XX Claim 2; Page 43-44; 58pp; English.
XX
CC The present sequence is the coding sequence for the corn cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a contig comprising
CC cDNA from corn roots and ear leaf sheath for sequences similar to those
CC encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and
CC Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote
CC or inhibit cell division and growth. The coding sequence and the protein
CC it encodes are useful in the production of transgenic plants which
CC produce increased or decreased amounts of the CDKI protein, in the
CC identification of herbicides, in genetic and physical mapping and in the
CC isolation of the CDKI gene in other organisms
XX
SQ Sequence 723 BP; 160 A; 202 C; 210 G; 151 T; 0 U; 0 Other;
Query Match 23.5%; Score 322; DB 3; Length 723;
Best Local Similarity 100.0%; Pred. No. 9.5e-100;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 636 GAGAGACGAGCCATCGAGCGGCGGCACGCGGAGCTCAGCGATCTGGATCGGATCTGG 695
Db |||||||
QY 696 CGGGGCACAAAGATGGCCCTCGCTACCGCGGCAACCGCGGTGCGAGCTGATCGTGC 755
Db |||||||
QY 172 CGGGGCACAAAGATGGCCCTCGCTACCGCGGCAACCGCGGTGCGAGCTGATCGTGC 231
QY 756 CGCAGACACAGATCAGAGTCTTCCGCCCGCGGCGGCCCGCCAGCCCAAGCGCT 815
Db |||||||
QY 232 CGCAGACACAGATCAGAGTCTTCCGCCCGCGGCGGCCCGCCAGCCCAAGCGCT 291
QY 816 TTGCTTCCAAAGTACAACTTCGATTCGTCGCGCGCTGCCCTCGAGCGCGGCGCGCT 875
Db |||||||
QY 292 TTGCTTCCAAAGTACAACTTCGATTCGTCGCGCGCTGCCCTCGAGCGCGGCGCGCT 351
QY 876 TCGAGTGGCGCGGTGGTTCAGATCTGAAGCGAGCGTTCGTCGCGTGCAGGTGAAGCT 935
Db |||||||
QY 352 TCGAGTGGCGCGGTGGTTCAGATCTGAAGCGAGCGTTCGTCGCGTGCAGGTGAAGCT 411
QY 936 AGAAGAGAAAGATGCCCCC 957
Db |||||||
QY 412 AGAAGAGAAAGATGCCCCC 433
RESULT 3
AAL02569
ID AAL02569 standard; cDNA; 149 BP.
XX
AC AAL02569;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen cDNA SEQ ID NO: 2570.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
XX
OS Homo sapiens.
XX
FN WO200155320-A2.
XX
PD 02-AUG-2001.
```

```
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
```


CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for preventing,
CC treating or ameliorating a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
CC can also be used as a food additive or preservative to increase or
CC decrease storage capabilities. The polynucleotides are useful for
CC chromosome identification. They are also useful as probes for diagnosing
CC a disorder related to the female reproductive system, particularly breast
CC and/or ovary cancer. They are also useful in the gene therapy of breast
CC and ovarian cancer. Nucleic acids, protein, antibodies, agonists and
CC antagonists from the present invention are useful in the diagnosis,
CC treatment and prevention of cancer, immune disorders, cardiovascular
CC disorders, wound healing, neurological diseases and infectious disease.
CC AAC59957 to AAC59965 and AAC54772 represents sequence used in the
CC exemplification of the present invention
XX SQ Sequence 1316 BP; 377 A; 287 C; 273 G; 372 T; 0 U; 7 Other;

Query Match 7.9%; Score 108; DB 3; Length 1316;
Best Local Similarity 100.0%; Pred. No. 2e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 1094 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAAGCTTACGT 1153
QY 1322 ACGGTGTCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTC 1369
Db 1154 ACGGTGTCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTC 1201

RESULT 5
AAC59034
ID AAC59034 standard; cDNA; 1557 BP.
XX AAC59034;
XX
XX 09-MAR-2001 (first entry)
XX
XX Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:262.
XX
XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
XX diagnosis; identification; cytostatic; neuroprotective; neurotropic;
XX immunomodulatory; relaxant; contraceptive; gynaecological;
XX antiinflammatory; cardiant; gene therapy; chromosome mapping;
XX linkage analysis; tissue identification; tissue typing; forensic; neural;
XX immune system; muscular; reproductive; gastrointestinal; pulmonary;
XX cardiovascular; renal; proliferative; ss.
XX
XX Homo sapiens.
XX
XX WO20005320-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US0005989.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-579444/54.
XX P-PSDB; AAB54269.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
XX treating, or ameliorating a medical condition, particular pancreatic
XX cancer, or for use in assays for diagnosing a pathological condition.
XX

PS
XX Claim 1; Page 701; 1379pp; English.
CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention
XX SQ Sequence 1557 BP; 412 A; 345 C; 315 G; 480 T; 0 U; 5 Other;

Query Match 7.9%; Score 108; DB 3; Length 1557;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 1211 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAAGCTTACGT 1270
QY 1322 ACGGTGTCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTC 1369
Db 1271 ACGGTGTCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTC 1318

RESULT 6
AAS00830
ID AAS00830 standard; cDNA; 1557 BP.
XX AAS00830;
XX
XX 04-JUL-2001 (first entry)
XX
XX Human cDNA clone HISCCL9 encoding cancer related protein 4.
XX
XX Human; cancer related protein; HISCCL9; food additive; preservative;
XX immunogen; antibody; bone cancer; adrenal cancer; bone marrow cancer;
XX breast cancer; gastrointestinal cancer; liver cancer; lung cancer;
XX urogenital cancer; immune disorder; Addison's disease; allergy;
XX autoimmune haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus;
XX Crohn's disease; multiple sclerosis; rheumatoid arthritis;
XX ulcerative colitis; acquired immunodeficiency syndrome; AIDS;
XX cardiovascular disorder; myocardial ischaemia; wound healing;
XX neurological disorder; Parkinson's disease; Alzheimer's disease;
XX cerebral anoxia; epilepsy; viral infection; bacterial infection;
XX fungal infection; parasitic infection; agonist; antagonist; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 259..667
XX FT /*tag= a
XX FT /product= "Cancer related protein 4"
XX FT 259..319
XX FT /*tag= b
XX FT 320..664
XX FT /*tag= c
XX FT /label= Mature_Cancer_related_protein_4

XX KW Human; cancer related protein; H15BL55; food additive; preservative;
KW immunogen; antibody; bone cancer; adrenal cancer; bone marrow cancer;
KW breast cancer; gastrointestinal cancer; liver cancer; lung cancer;
KW urogenital cancer; immune disorder; Addison's disease; allergy;
KW autoimmune haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus;
KW Crohn's disease; multiple sclerosis; rheumatoid arthritis;
KW ulcerative colitis; acquired immunodeficiency syndrome; AIDS;
KW cardiovascular disorder; myocardial ischaemia; wound healing;
KW neurological disorder; Parkinson's disease; Alzheimer's disease;
KW cerebral anoxia; epilepsy; viral infection; bacterial infection;
KW fungal infection; parasitic infection; agonist; antagonist; ss.
XX OS Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
XX CDS 342..482
FT /*tag= a
FT /product= "Cancer related protein 21"
XX WO200118014-A1.
XX PN
XX PD 15-MAR-2001.
XX PF 30-AUG-2000; 2000WO-US023794.
XX PR 03-SEP-1999; 99US-0152296P.
XX PR 06-OCT-1999; 99US-0158003P.
XX XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI
XX PI Roschke V;
XX DR
XX DR WPI; 2001-235186/24.
XX DR P-PSDB; AAU00886.
XX XX
XX PT Twenty nine nucleic acid molecules encoding human cancer associated
XX PT proteins, useful in the prevention, treatment and diagnosis of cancer,
XX PT immune disorders, cardiovascular disorders and neurological diseases.
XX PS
XX PS Disclosure; Page 379-380; 427pp; English.
XX CC The sequence encodes a novel Human cancer related protein. The
XX CC polynucleotides and polypeptides are useful for preventing, treating or
XX CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
XX CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
XX CC as a food additive or preservative to increase or decrease storage
XX CC capabilities. The polynucleotide are useful for chromosome
XX CC identification. The nucleic acids, protein, antibodies, agonists and
XX CC antagonists are useful in the diagnosis, treatment and prevention of
XX CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
XX CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
XX CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
XX CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
XX CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
XX CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
XX CC myocardial ischaemias, wound healing, neurological diseases (e.g.
XX CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
XX CC and infectious diseases such as viral, bacterial, fungal and parasitic
XX CC infections. Numerous examples of each type of disorder are given in the
XX CC specification
XX SQ Sequence 2116 BP; 606 A; 485 C; 390 G; 629 T; 0 U; 6 Other;
Query Match 7.9%; Score 108; DB 5; Length 2116;
Best Local Similarity 100.0%; Pred. No. 1.8e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAGCGCGCGCTCTAGAGATCCAGCTTACGT 1321
DB 1977 AAAAAAAAAAAAAAAAAAAGCGCGCGCTCTAGAGATCCAGCTTACGT 2036
QY 1322 ACGCGTGCATCGACGTCTCTCTATAGTGTACCTAAATTCA 1369

DB 2037 ACGCGTGCATCGACGTCTCTCTATAGTGTACCTAAATTCA 2084
RESULT 9
ID AAH33453 standard; cDNA; 701 BP.
XX AAH33453;
AC AAH33453;
XX 03-SEP-2001 (first entry)
XX Human colon cancer antigen encoding cDNA SEQ ID NO:509.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX Homo sapiens.
XX WO200122920-A2.
XX PD 05-APR-2001.
XX PF 28-SEP-2000; 2000WO-US026524.
XX PR 29-SEP-1999; 99US-0157137P.
XX PR 03-NOV-1999; 99US-0163280P.
XX XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX DR WPI; 2001-235357/24.
XX DR P-PSDB; AAG74022.
XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX PS Claim 1; Page 2588; 9803pp; English.
XX CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
XX CC proteins are collectively known as colon cancer antigens. The colon
XX CC cancer antigens have cytostatic activity and can be used in gene therapy
XX CC and vaccine production. N and P may be used in the prevention, diagnosis
XX CC and treatment of diseases associated with inappropriate P expression. For
XX CC example, N and P may be used to treat disorders associated with decreased
XX CC expression by rectifying mutations or deletions in a patient's genome
XX CC that affect the activity of P by expressing inactive proteins or to
XX CC supplement the patient's own production of P. Additionally, N may be used
XX CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
XX CC into a host cell and culturing the cell to express the proteins. N and P
XX CC can be used in the prevention, diagnosis and treatment of colorectal
XX CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
XX CC sequences used in the exemplification of the present invention. N.B.
XX CC pages 666 to 682 and page 7053 of the sequence listing were missing at
XX CC time of publication, meaning no sequences are present for SEQ ID NO:1027
XX CC to 1052, 7921 and 7922
XX SQ Sequence 701 BP; 200 A; 164 C; 146 G; 181 T; 0 U; 10 Other;
Query Match 7.7%; Score 106; DB 4; Length 701;
Best Local Similarity 100.0%; Pred. No. 1.1e-26;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1264 AAAAAAAAAAAAAAAAAAAGCGCGCGCTCTAGAGATCCAGCTTACGT 1323
DB 459 AAAAAAAAAAAAAAAAAAAGCGCGCGCTCTAGAGATCCAGCTTACGT 518
QY 1324 GCGTGCATCGACGTCTCTCTATAGTGTACCTAAATTCA 1369
DB 519 GCGTGCATCGACGTCTCTCTATAGTGTACCTAAATTCA 564

RESULT 10

ABK69104
ID ABK69104 standard; cDNA; 1928 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

RESULT 11

AAD05156

ID AAD05156 standard; cDNA; 2152 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAGCTTAGCT 1321

1822 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAGCTTAGCT 1881

1322 ACGGTGCGATCGGAGCGTCATAGCTCTTCTATAGTGCACCTAAAT 1366

1882 ACGGTGCGATCGGAGCGTCATAGCTCTTCTATAGTGCACCTAAAT 1926

17-JUL-2001 (first entry)

Human secreted protein-encoding gene 5 cDNA clone HGCAC66, SEQ ID NO:46.

Human; secreted protein; proliferative disorder; cancer; tumour;

foetal abnormality; developmental abnormality; haematopoietic disorder;

immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

inflammation; allergy; neurological disorder; Alzheimer's disease;

Parkinson's disease; cognitive disorder; schizophrenia; asthma;

skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

cardiovascular disorder; angiogenic disorder; kidney disorder;

gastrointestinal disorder; pregnancy-related disorder;

endocrine disorder; infection; wound healing; vulnerability; cell culture;

chemotaxis; food additive; gene therapy; binding partner identification;

ss.

Homo sapiens.

Key

Location/Qualifiers

78..173

/*tag= a

/*product= "Human secreted protein"

78..140

/*tag= b

141..170

/*tag= c

/*product= "Human mature secreted protein"

WO200134769-A2.

17-MAY-2001.

01-NOV-2000; 2000WO-US030040.

05-NOV-1999; 99US-0163580P.

30-JUN-2000; 2000US-0215130P.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;

P-PSDB; AAE01267.

New isolated nucleic acid molecule encoding a human secreted protein is

used in preventing, treating or ameliorating a medical condition.

Claim 1; Page 427-428; 519pp; English.

AAD05121-AAD05203 represent cDNAs corresponding to 24 human secreted

protein genes, and AAE01232-AAE01311 represent the proteins they encode.

AAE01312-AAE01340 represent human secreted protein variants or fragments.

The secreted proteins and their genes are useful for preventing, treating

or ameliorating medical conditions, e.g., by protein or gene therapy.

Pathological conditions can be diagnosed by determining the amount of the

new protein in a sample or by determining the presence of mutations in

the new genes. Specific uses are described for each of the 24 genes,

based on the tissues in which they are most highly expressed, and include

CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein-encoding cDNA of the
CC invention
XX
SQ Sequence 2152 BP; 605 A; 465 C; 407 G; 664 T; 0 U; 11 Other;

Query Match 7.6%; Score 104; DB 4; Length 2152;
Best Local Similarity 100.0%; Pred. No. 3.9e-26;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1266 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGTCCAAAGCTTAGTCAGC 1325
DB 1869 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGTCCAAAGCTTAGTCAGC 1928

QY 1326 GTGCATCGGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1369
DB 1929 GTGCATCGGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1972

RESULT 12
ABQ54253/c
ID ABQ54253 standard; cDNA; 845 BP.
XX
AC ABQ54253;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HCCMD55 cDNA, SEQ ID NO:133.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytotoxic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.

OS Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

DR P-PSDB; ABP41176.

XX
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ

Isolated nucleic acid molecules encoding novel ovarian polypeptides,
useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
cancer), immune disorders, cardiovascular disorders and neurological
diseases.

Claim 1; SEQ ID NO 133; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-
ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
encompasses polypeptides 90% identical and polynucleotides 95% identical
to the sequences of the invention. The invention additionally relates to
recombinant vectors and host cells comprising human ovarian antigen
polynucleotides, antibodies against human ovarian antigens, and the use
of ovarian antigen polynucleotides and polypeptides in diagnosing,
treating, prognosing or preventing various ovary and/or breast-related
disorders. Such conditions include ovarian cancer and breast cancer, and
metastatic tumours of ovarian or breast origin, reproductive system
disorders (e.g., infertility, disorders of pregnancy, anovulation,
polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
vaginitis), immune disorders (e.g., congenital and acquired
immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
blood-related disorders (e.g., anaemia), cardiovascular disorders,
respiratory disorders, neurological disorders, gastrointestinal disorders
and urinary system disorders. Ovarian antigen polypeptides and
polynucleotides may also be used in screening for compounds which
modulate ovarian antigen expression or activity. The polynucleotides may
further be used for gene therapy, chromosome mapping, in the
identification of individuals and in forensic analysis, and the
polypeptides may be used as food additives or to prepare antibodies
useful in disease diagnosis, drug targeting and phenotyping. The present
sequence represents cDNA encoding a human ovarian antigen of the
invention. Note: The sequence data for this patent did not form part of
the printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pt_sequences

Sequence 845 BP; 182 A; 226 C; 234 G; 199 T; 0 U; 4 Other;

Query Match 7.5%; Score 103; DB 6; Length 845;

Best Local Similarity 100.0%; Pred. No. 1.1e-25;

Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1267 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGTCCAAAGCTTAGTCAGC 1326
DB 178 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGTCCAAAGCTTAGTCAGC 119

QY 1327 TGCATCGGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1369

DB 118 TGCATCGGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 76

RESULT 13

AAD13354

ID AAD13354 standard; cDNA; 1088 BP.

XX AAD13354;

XX 16-OCT-2001 (first entry)

XX Human secreted protein-encoding gene 10 cDNA clone HTGF16, SEQ ID NO:20.

KW Human secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery; cell culture;
KW chemotaxis; food additive; gene therapy; binding partner identification;

ss.
Homo sapiens.
Key Location/Qualifiers
CDS 160..477
/tag= a
/product= "Human secreted protein precursor"
sig_peptide 160..282
/tag= b
mat_peptide 283..474
/tag= c
/product= "Mature human secreted protein"
WO200154708-A1.
02-AUG-2001.
17-JAN-2001; 2001WO-US001434.
31-JAN-2000; 2000US-0179065P.
04-FEB-2000; 2000US-0180628P.
18-AUG-2000; 2000US-0226279P.
05-DEC-2000; 2000US-0251988P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Komatsculis GA, Baker KP, Birse CE, Soppet DR;
Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
Fiscella M, Ni J, Ruben SM, Barash SC;
WPI; 2001-488743/53.
P-PSDB; AAE07060.
New isolated nucleic acids and polypeptides, useful for diagnosing,
treating and/or preventing human diseases and disorders.
Claim 1; Page 452-453; 558pp; English.
AAD13345-AAD13401 represent cDNAs corresponding to 22 human secreted
protein genes, and AAE07051-AAE07105 represent the proteins they encode.
AAE07106-AAE07129 represent human secreted protein fragments or variants.
The genes and their secreted proteins are useful for preventing, treating
or ameliorating medical conditions, e.g., by protein or gene therapy.
Pathological conditions can be diagnosed by determining the amount of the
new protein in a sample or by determining the presence of mutations in
the new genes. Specific uses are described for each of the 22 genes,
based on the tissues in which they are most highly expressed, and include
developing products for the diagnosis or treatment of proliferative
disorders, cancer, tumours, foetal and developmental abnormalities,
haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
diseases (e.g., rheumatoid arthritis), inflammation, allergies, and
neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
angiogenic disorders, kidney disorders, gastrointestinal disorders,
pregnancy-related disorders, endocrine disorders, and infections. The
proteins can also be used to aid wound healing and epithelial cell
proliferation, to prevent skin aging due to sunburn, to maintain organs
before transplantation, for supporting cell culture of primary tissues,
to regenerate tissues, to identify their cognate ligands or binding
partners, and in chemotaxis, and can be used as a food additive or
preservative to modify storage properties. Antibodies specific for a
protein of the invention can be used in alleviating symptoms associated
with the disorders mentioned above, and in diagnostic immunoassays e.g.,
radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
present sequence represents a human secreted protein-encoding cDNA of the
invention.
Sequence 1088 BP; 284 A; 246 C; 284 G; 269 T; 0 U; 5 Other;
Query Match 7.5%; Score 103; DB 4; Length 1088;

Best Local Similarity 100.0%; Pred. No. 1e-25;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAGCTTACGT 1321
Db 986 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAGCTTACGT 1045
Qy 1322 ACGGTCGATCGGACGTCATAGCTCTTCTATAGTGTACCTAA 1364
Db 1046 ACGGTCGATCGGACGTCATAGCTCTTCTATAGTGTACCTAA 1088
RESULT 14
ABL90818/c
ID ABL90818 standard; cDNA: 2049 BP.
XX
AC ABL90818;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1380.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
XX
PT P-PSDB; ABB90409.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders.
PS Claim 4; SEQ ID NO 1380; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2049 BP; 383 A; 571 C; 686 G; 400 T; 0 U; 9 Other;
Query Match 7.2%; Score 99; DB 6; Length 2049;

Best Local Similarity 100.0%; Pred. No. 2e-24;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1271 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTACGCGTGCA 1330
DB 321 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTACGCGTGCA 262

QY 1331 TGGACGCTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
DB 261 TGGACGCTCATAGCTCTTCTATAGTGTACCTAAATTCA 223

RESULT 15
ID ABZ5039
XX ABZ5039 standard; cDNA; 570 BP.
XX AC ABZ5039;
XX DT 28-MAR-2003 (first entry)
XX XX Aspergillus oryzae polynucleotide SEQ ID NO 4152.
XX DE Aspergillus oryzae; fermentation; fungus; industrial; EST;
XX KW expressed sequence tag; gene; ss.
XX XX Aspergillus oryzae.
XX OS
XX PN WO200279476-A1.
XX PD 10-OCT-2002.
XX PF 22-MAR-2002; 2002WO-IB000890.
XX PR 30-MAR-2001; 2001JP-00098371.
XX XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PA (NARE-) NAT RES INST BREWING.
XX PA (NORQ) NAT FOOD RES INST MIN AGRIC.
XX PI Machida M, Akita O, Kaehiwagi Y, Kitamoto K, Horiuchi H;
XX PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX DR WPI; 2003-046817/04.
XX PT Detection of expression of specific Aspergillus genes for monitoring the
XX PT fermentation and growth conditions of the fungus, using DNA probes.
XX PS Claim 1; SEQ ID NO 4152; 48pp + Sequence Listing; Japanese.
XX CC The invention relates to a polynucleotide having any of 6006 specific
XX CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
XX CC specific culture conditions including one or more of eutrophic, low
XX CC oligotrophic, solid, early germination, alkaline, high temperature, low
XX CC temperature or maltose culture or polynucleotides stringently hybridising
XX CC to these sequences. The polynucleotides are useful for monitoring the
XX CC progress of fermentation and the growth conditions of a fungus,
XX CC especially of Aspergillus oryzae which is widely used in industrial
XX CC fermentation. Also monitoring for fungal contamination. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences

QY 1272 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTACGCGTGCA 1331
DB 286 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTACGCGTGCA 345

QY 1332 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369

Db 346 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 383

Search completed: October 2, 2004, 06:11:50
Job time : 567 secs

Geisler R., Wilson S.W., de Hertog J., Kimata K. and Zivkovic D.
 Heparan Sulfate 6-O-Sulfotransferase Is Essential for Muscle
 Development in Zebrafish
 J. Biol. Chem. 278 (33), 31118-31127 (2003)
 12782624
 REFERENCE 2 (bases 1 to 2658)
 AUTHORS Bink R.J.
 TITLE Direct Submission
 JOURNAL Submitted (19-AUG-2002) Bink R.J., Netherlands Inst. for Dev.
 Biology, Hubrecht Laboratory, Uppsalalaan 8, 3584CT Utrecht,
 NETHERLANDS

FEATURES
 source
 1. .2658
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /chromosome="14"
 1. .2658
 /gene="hs6st"
 314..1720
 /gene="hs6st"
 /function="brain and fin development, myogenesis"
 /codon_start=1
 /product="heparan sulfate 6-O-sulfotransferase"
 /protein_id="CAD44529.1"
 /db_xref="GI:29150001"
 /db_xref="SPTREMBL:Q600H9"
 /translation="MDGKSYSLRLIALMLLFFGGVLYQICSTSDWQLHLASLS
 RLGSRAPDRLNGAGDPYSSDGALVRFVFNFTTKDLSRADFHIGDDVIVFL
 HIQTGGTTFGRHLVNIQLRCPCECHAGQKCTCYRGRDITWLFKSTGWSGLH
 ADMELTNVPSFMSNRESQRRMTPSRNYIITLRDPVMRYLSEWHVQRGATKA
 SKMCDRLPTLTELPCYDGDWGSGLREPMVCPYVNNRQTRMLADLSLVCYN
 LTVSNQRWAMLLSAKRLRNWAFGLTEYORKTOYLFPHTRLSFIAPFTOLNGT
 RAASVEVEPTQRIKRLNQMDVELYEARLFLQRPQARQRRRQRRIQRK
 LRKVKSWLGVPTGKAVFKPTKEPTMTQSPAFBEKQADAEQTLESETEGQVEENWLE
 EDGEIMLDYSENVEQWR"

ORIGIN
 Query Match 7.9%; Score 108; DB 5; Length 2658;
 Best Local Similarity 100.0%; Pred. No. 3.8e-44;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
 Db 2531 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGT 2590

Qy 1322 ACGCGTCATGCGAGCTCATGCTCTCTATAGTGTACCTAAATTC 1369
 Db 2591 ACGCGTCATGCGAGCTCATGCTCTCTATAGTGTACCTAAATTC 2638

RESULT 2
 BD275987
 LOCUS 62 Human secreted proteins.
 DEFINITION
 ACCESSION BD275987
 VERSION BD275987.1 GI:33085755
 KEYWORDS JP 2002543771-A/65.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2657)
 Birse,C.E., Moutet,P.A., Florence,K.A., Ruben,S.M.,
 Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
 Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
 Patent: JP 2002543771-A 65 24-DEC-2002;
 Human Genome Sciences Inc
 OS Homo sapiens
 PN JP 2002543771-A/65
 PD 24-DEC-2002
 PR 06-APR-2000 JP 2000611564

BD275987 2657 bp DNA linear PAT 17-JUL-2003
 62 Human secreted proteins.
 DEFINITION
 ACCESSION BD275987
 VERSION BD275987.1 GI:33085755
 KEYWORDS JP 2002543771-A/65.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2657)
 Birse,C.E., Moutet,P.A., Florence,K.A., Ruben,S.M.,
 Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
 Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
 Patent: JP 2002543771-A 65 24-DEC-2002;
 Human Genome Sciences Inc
 OS Homo sapiens
 PN JP 2002543771-A/65
 PD 24-DEC-2002
 PR 06-APR-2000 JP 2000611564

PR 09-APR-1999 US 60/128693,26-APR-1999 US 60/130991 PI
 Charles e birse,paul a mouret,kimberly a florence,steven m PI
 ruben,
 PI george a komatsoulis,jian ni,reinhard ebner,david w PI
 lafleur,henrik s olsen,
 PI yanggu shi,daniel r soppet,craig a rosen,paul e young CC
 FH key Location/Qualifiers
 FT SITE (179)
 FT /note='n equals a,t,g, or c'.
 Location/Qualifiers
 1. .2657
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source
 1. .2657
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 7.1%; Score 98; DB 6; Length 2657;
 Best Local Similarity 100.0%; Pred. No. 5.3e-39;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1272 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGGTGCAT 1331
 Db 2307 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGGTGCAT 2366

Qy 1332 GCGACGTCATAGCTCTCTTATAGTGTACCTAAATTC 1369
 Db 2367 GCGACGTCATAGCTCTCTTATAGTGTACCTAAATTC 2404

RESULT 3
 BD276004 2454 bp DNA linear PAT 17-JUL-2003
 LOCUS 62 Human secreted proteins.
 DEFINITION
 ACCESSION BD276004
 VERSION BD276004.1 GI:33085772
 KEYWORDS JP 2002543771-A/82.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2454)
 Birse,C.E., Moutet,P.A., Florence,K.A., Ruben,S.M.,
 Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
 Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
 Patent: JP 2002543771-A 82 24-DEC-2002;
 Human Genome Sciences Inc
 OS Homo sapiens
 PN JP 2002543771-A/82
 PD 24-DEC-2002
 PR 06-APR-2000 JP 2000611564
 PR 09-APR-1999 US 60/128693,26-APR-1999 US 60/130991 PI
 Charles e birse,paul a mouret,kimberly a florence,steven m PI
 ruben,
 PI george a komatsoulis,jian ni,reinhard ebner,david w PI
 lafleur,henrik s olsen,
 PI yanggu shi,daniel r soppet,craig a rosen,paul e young CC
 FH key Location/Qualifiers
 FT SITE (2317)
 FT /note='n equals a,t,g, or c'.
 Location/Qualifiers
 1. .2454
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source
 1. .2454
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 7.1%; Score 97; DB 6; Length 2454;
 Best Local Similarity 100.0%; Pred. No. 1.7e-38;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1273 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGGTGCATG 1332

Db 2105 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATG 2164

Qy 1333 CGAGTCATAGCTCTTCTATAGTGTACCTACCTAAATTC A 1369
|||||

Db 2165 CGAGTCATAGCTCTTCTATAGTGTACCTAAATTC A 2201
|||||

RESULT 4

LOCUS AR137827 1255 bp DNA linear PAT 16-JUN-2001

DEFINITION Sequence 31 from patent US 6197561.

ACCESSION AR137827

VERSION AR137827.1 GI:14479336

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1255)

AUTHORS Martino-Catt,S.J., Wang,H., Beach,L.R., Wang,X. and Bowen,B.A.

TITLE Genes controlling phytate metabolism in plants and uses thereof

JOURNAL Patent: US 6197561-A 31 06-MAR-2001;

FEATURES
Location/Qualifiers
source 1..1255
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATGCG 1334
|||||

Db 1161 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATGCG 1220
|||||

Qy 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTC A 1369
|||||

Db 1221 ACGTCATAGCTCTTCTATAGTGTACCTAAATTC A 1255
|||||

RESULT 5

LOCUS AR170186 1255 bp DNA linear PAT 17-DEC-2001

DEFINITION Sequence 31 from patent US 6291224.

ACCESSION AR170186

VERSION AR170186.1 GI:17908145

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1255)

AUTHORS Martino-Catt,S.J., Wang,H., Beach,L.R., Wang,X. and Bowen,B.A.

TITLE Genes controlling phytate metabolism in plants and uses thereof

JOURNAL Patent: US 6291224-A 31 18-SEP-2001;

FEATURES
Location/Qualifiers
source 1..1255
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATGCG 1334
|||||

Db 1161 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATGCG 1220
|||||

Qy 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTC A 1369
|||||

Db 1221 ACGTCATAGCTCTTCTATAGTGTACCTAAATTC A 1255
|||||

Db 2105 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATG 2164

Qy 1333 CGAGTCATAGCTCTTCTATAGTGTACCTACCTAAATTC A 1369
|||||

Db 2165 CGAGTCATAGCTCTTCTATAGTGTACCTAAATTC A 2201
|||||

RESULT 6

LOCUS BD247539 1450 bp DNA linear PAT 17-JUL-2003

DEFINITION Methods for transforming plastids.

ACCESSION BD247539

VERSION BD247539.1 GI:33057309

KEYWORDS JP 2002531096-A/3.

SOURCE Brassica sp.

ORGANISM Brassica sp.

REFERENCE 1 (bases 1 to 1450)

AUTHORS Chaudhuri,S.

TITLE Methods for transforming plastids

JOURNAL Patent: JP 2002531096-A 3 24-SEP-2002;

COMMENT CALGENE LLC

OS Brassica sp. (mustard)

PN JP 2002531096-A/3

PD 24-SEP-2002

PF 24-NOV-1999 JP 2000585430

PR 25-NOV-1998 US 60/109892

PI SUMITA CHAUDHURI

PC C12N15/09,A01H5/00,C12N5/10,C12N5/00,C12N5/00 CC Methods for transforming plastids

PH Key Location/Qualifiers

FT source 1..1450
/organism="Brassica sp. (mustard)"

FEATURES
Location/Qualifiers
source 1..1450
/organism="Brassica sp."
/mol_type="genomic DNA"
/db_xref="taxon:3717"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1450;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATGCG 1334
|||||

Db 1346 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATGCG 1405
|||||

Qy 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTC A 1369
|||||

Db 1406 ACGTCATAGCTCTTCTATAGTGTACCTAAATTC A 1440
|||||

RESULT 7

LOCUS AR066494 1581 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5850022.

ACCESSION AR066494

VERSION AR066494.1 GI:5996710

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1581)

AUTHORS Dehesh,K., Voelker,T. and Hawkins,D.

TITLE Production of myristate in plant cells

JOURNAL Patent: US 5850022-A 1 15-DEC-1998;

FEATURES
Location/Qualifiers
source 1..1581
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGTACGCGTCATGCG 1334
|||||

Db 1479 AAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTCAGCGTCATCGG 1538
QY 1335 ACCTCATAGCTCTCTATAGTGTACCTAAATTC 1369
Db 1539 ACCTCATAGCTCTCTATAGTGTACCTAAATTC 1573

RESULT 8
LOCUS I60018 1581 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5654495.
ACCESSION I60018
VERSION I60018.1 GI:2478650
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1581)
AUTHORS Voelker, I. Alois. and Davies, H. Maelor.
TITLE Production of myristate in plant cells
JOURNAL Patent: US 5654495-A 1 05-AUG-1997;
FEATURES Location/Qualifiers
source 1..1581
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 6.9%; Score 95; DB 6; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1275 AAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTCAGCGTCATCGG 1334
Db 1479 AAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTCAGCGTCATCGG 1538
QY 1335 ACCTCATAGCTCTCTATAGTGTACCTAAATTC 1369
Db 1539 ACCTCATAGCTCTCTATAGTGTACCTAAATTC 1573

RESULT 9
LOCUS BD136416 2015 bp DNA linear PAT 18-SEP-2002
DEFINITION 95 human secretory proteins.
ACCESSION BD136416
VERSION BD136416.1 GI:23231361
KEYWORDS JP 2002506627-A/103.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2015)
AUTHORS Ruben, S.M., Ni, J., Rosen, C.A., Yu, G.L., Young, P.E., Fen, P., Soppet, D.R., Wei, Y.F., Endress, G.A., Duan, R.D., Kyaw, H., Ebner, R., Lafleur, D.W., Olsen, H.S., Shi, Y. and Moore, P.A.
TITLE 95 human secretory proteins
JOURNAL Patent: JP 2002506627-A 103 05-MAR-2002;
COMMENT HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002506627-A/103
PD 05-MAR-2002
PF 18-MAR-1999 JP 2000536733
PR 19-MAR-1998 US 60/078566, 19-MAR-1998 US 60/078576 PR
19-MAR-1998 US 60/078573, 19-MAR-1998 US 60/078574 PR
19-MAR-1998 US 60/078579, 19-MAR-1998 US 60/078578 PR
19-MAR-1998 US 60/078581, 19-MAR-1998 US 60/078577 PR
19-MAR-1998 US 60/078563, 01-APR-1998 US 60/080314 PR
01-APR-1998 US 60/080312, 01-APR-1998 US 60/080313 PI STEVEN M RUBEN, JUAN NI, CRAIG A ROSEN, GUO LIANG YU, PAUL E YOUNG, LIANG YU, PAUL E YOUNG, PI PING FENG, PI DANIEL R Soppet, YING FEI WEI, GREGORY A ENDRESS, ROXANNE D DUAN,

PI HLA KYAW,
PI REINHARD EBNER, DAVID W LAFLEUR, HENRIK S OLSEN, YANGGU SHI, PAUL
PI A MOORE
PC C12N15/09, A61K38/00, A61K48/00, A61P43/00, C07K14/47, C07K16/18,
PC C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/53, G01N33/PC
566,
PC C12N15/00, A61K37/02, C12N5/00
CC n equals a.t.g, or c
CC n equals a.t.g, or c
CC n equals a.t.g, or c
CC n equals a.t.g, or c
CC n equals a.t.g, or c
CC n equals a.t.g, or c
CC n equals a.t.g, or c
FH Key
FT source 1..2015
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 6.9%; Score 95; DB 6; Length 2015;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTCAGCGTCATCGT 1321
Db 1548 AAAAAAAAAAAGGCGCGCTCTAGAGGATCAAGCTTACGTCAGCGTCATCGT 1607
QY 1322 ACAGCGTCATCGGTCATAGCTTCTCTATAGTG 1356
Db 1608 ACAGCGTCATCGGTCATAGCTTCTCTATAGTG 1642

RESULT 10
LOCUS AX301825 2486 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 4 from Patent WO0185768.
ACCESSION AX301825
VERSION AX301825.1 GI:17382883
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Gu, W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 4 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1..2486
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13077.1"
/db_xref="GI:17382884"
/db_xref="REMBEML:CAD13077"
/translation="NTHYRESWACRYSGIPGSHASVERSGLSLPAHPASLAAL
AASNTAGSKLELDTFQSSQALDLSNNAIRSHPAFSTLSLHSLVLDLTDLNLT
LPLAGLGGMLHKLGNLALQAFSKDFPKRLILEVPIAYOCCPYGMCASFCKASGQ
WEAPDLHLDDESSKRPLGLARQENHYDQDLDELQEMEDSKPHSPVQCSPTGPP
KPCYILFSGWIGLRAVAILLVLSVLCNGLVITVFPAGGAPLPVPKVVVGAAGANTL
TGISGGLIASVDALTFQGFSEYGARWETGLGRATGLAVLGSEASVLLTLAAVQCS
VSVSVRAYGKSPSGSVRAGVLGCLALAGLAALPLASVGEYASPLCLPYAPPEGQ
PAAIGFTVALVMNSFCFLVAVAGAIKLYCDLPRGDFEAVMDCMVRHVWLIIFADGL

LYCFVAFLSFASMLGLFPVTPPEAVKSVLLVLPACLNPLLLXLLFNPRDDIRLR
PRAGDSGLAYAAAGLEKSSCDSTQALVAFSDVLDLLEASGRPPGLTYGPPSVT
LISCOOPGAPRLRGSHCVPEGNHFGNFPQPSMDGELLRLRAEGSTPAGGSLGGGGPQ
SGLAFASHV"

ORIGIN
Query Match 6.9%; Score 95; DB 6; Length 2486;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAGCGCGCTCTAGAGTCAAGCTTACGCGTGCATGCG 1334
|||||
Db 2389 AAAAAAAAAAAGCGCGCTCTAGAGTCAAGCTTACGCGTGCATGCG 2448
|||||

QY 1335 ACCTCATAGCTCTCTATAGTGTCACTAAATCA 1369
|||||
Db 2449 ACCTCATAGCTCTCTATAGTGTCACTAAATCA 2483
|||||

RESULT 11
AR184178
LOCUS AR184178 944 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 122 from patent US 6342581.
ACCESSION AR184178
VERSION AR184178.1 GI:20228147
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 944)
AUTHORS Rosen, C.A., Ruben, S.M., Olsen, H.S. and Ebner, R.
TITLE Secreted protein HLHFP03
JOURNAL Patent: US 6342581-A 122 29-JAN-2002;
FEATURES
source
1..944
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 6.6%; Score 90; DB 6; Length 944;
Best Local Similarity 100.0%; Pred. No. 7.1e-35;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAGCGCGCTCTAGAGGATCAAGCTTACGT 1321
|||||
Db 835 AAAAAAAAAAAGCGCGCTCTAGAGGATCAAGCTTACGT 894
|||||
QY 1322 ACGCGTGATGCGAGCTCATAGCTCTTCTA 1351
|||||
Db 895 ACGCGTGATGCGAGCTCATAGCTCTTCTA 924
|||||

RESULT 12
BD192707
LOCUS BD192707 944 bp DNA linear PAT 17-JUL-2003
DEFINITION 123 human secreted proteins.
ACCESSION BD192707
VERSION BD192707.1 GI:33002446
KEYWORDS JP 2002513295-A/121.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 944)
AUTHORS Fischer, C.L., Rosen, C.A., Soppet, D.R., Ruben, S.M., Kyaw, H., Li, Y., Zeng, Z., Lafleur, D.W., Moore, P.A., Shi, Y., Ols, H.S., Ebner, R. and Brewer, L.A.

TITLE 123 human secreted proteins
JOURNAL Patent: JP 2002513295-A 121 08-MAY-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002513295-A/121
PD 08-MAY-2002

PF 07-JUL-1998 JP 1999508744
PR 08-JUL-1997 US 60/051926, 08-JUL-1997 US 60/051929 PR
08-JUL-1997 US 60/052793, 08-JUL-1997 US 60/051925 PR
08-JUL-1997 US 60/051931, 08-JUL-1997 US 60/051932 PR
08-JUL-1997 US 60/052803, 08-JUL-1997 US 60/052732 PR
08-JUL-1997 US 60/051916, 08-JUL-1997 US 60/051930 PR
08-JUL-1997 US 60/051918, 08-JUL-1997 US 60/051920 PR
08-JUL-1997 US 60/051919, 08-JUL-1997 US 60/051928 PR
08-JUL-1997 US 60/052795, 08-JUL-1997 US 60/052733 PR
18-AUG-1997 US 60/055948, 18-AUG-1997 US 60/055722 PR
18-AUG-1997 US 60/055723, 18-AUG-1997 US 60/055949 PR
18-AUG-1997 US 60/055964, 18-AUG-1997 US 60/055947 PR
18-AUG-1997 US 60/055950, 18-AUG-1997 US 60/055953 PR
18-AUG-1997 US 60/056360, 18-AUG-1997 US 60/056884 PR
18-AUG-1997 US 60/055984, 18-AUG-1997 US 60/055954 PR
12-SEP-1997 US 60/058785, 12-SEP-1997 US 60/058664 PR
12-SEP-1997 US 60/058660, 12-SEP-1997 US 60/058661 PI
L FISCHER, CRAIG A ROSEN, DANIEL R SOPPET, STEVEN M RUBEN, PI
KYAW, YI LI, ZHIZHEN ZENG, DAVID W LAFLEUR, PAUL A MOORE, YANGGU SHI, PI
HENRIK S OLSEN, REINHARD EBNER, LAURIE A BREWER
PC C07H21/04, C12N15/63
CC n equals a,t,g, or c
CC n equals a,t,g, or c
CC n equals a,t,g, or c
FH Key Location/Qualifiers
FT misc feature (932)
FT misc feature (942)
FT misc feature (944).

FEATURES
source
1..944
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 6.6%; Score 90; DB 6; Length 944;
Best Local Similarity 100.0%; Pred. No. 7.1e-35;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAGCGCGCTCTAGAGGATCAAGCTTACGT 1321
|||||
Db 835 AAAAAAAAAAAGCGCGCTCTAGAGGATCAAGCTTACGT 894
|||||
QY 1322 ACGCGTGATGCGAGCTCATAGCTCTTCTA 1351
|||||
Db 895 ACGCGTGATGCGAGCTCATAGCTCTTCTA 924
|||||

RESULT 13
BD275352
LOCUS BD275352 1270 bp DNA linear PAT 17-JUL-2003
DEFINITION 49 Human Secreted Proteins.
ACCESSION BD275352
VERSION BD275352.1 GI:33085120
KEYWORDS JP 2002539787-A/28.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1270)
AUTHORS Komatsoulis, G., Rosen, C.A. and Ruben, S.M.
TITLE 49 Human Secreted Proteins
JOURNAL Patent: JP 2002539787-A 28 26-NOV-2002;
Human Genome Sciences Inc
COMMENT OS Homo sapiens
PN JP 2002539787-A/28
PD 26-NOV-2002

PF 16-MAR-2000 JP 2000606742
PR 23-MAR-1999 US 60/126054, 10-DEC-1999 US 60/169916 PI
george komatsoulis, craig a rosen, steven m ruben CC
FH Key Location/Qualifiers

```

FT SITE (553)
/Note='n equals a,t,g, or c'
FT SITE (1070)
/Note='n equals a,t,g, or c'
FT SITE (1208)
/Note='n equals a,t,g, or c'
FEATURES
  Location/Qualifiers
    1..1270
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
ORIGIN
Query Match 6.5%; Score 89; DB 6; Length 1270;
Best Local Similarity 100.0%; Pred. No. 2.3e-34;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1270 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 981 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1040
QY 1330 ATGCGAGTCATAGCTCTCTTATAGTGC 1358
Db 1041 ATGCGAGTCATAGCTCTCTTATAGTGC 1069
ORIGIN
BD130531 2581 bp DNA linear PAT 18-SEP-2002
DEFINITION Ligand receptors and utilization thereof.
ACCESSION BD130531
VERSION BD130531.1 GI:23225476
KEYWORDS JP 2002501083-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2581)
AUTHORS Glucksmann,A.M. and Robison,K.
TITLE Ligand receptors and utilization thereof
JOURNAL Patent: JP 2002501083-A 1 15-JAN-2002;
MILLENNIUM PHARMACEUTICALS INC
COMMENT OS Unidentified
PN JP 2002501083-A/1
PD 15-JAN-2002
PF 20-JAN-1999 JP 2000528599
PR 26-JAN-1998 US 09/013634
PI ALEXANDRA M GLUCKSMANN,KEITH ROBISON
PC C07K14/705,C07K16/28,C12N5/10,C12N15/09,C12P21/02,C12Q1/68, PC
G01N33/53,
PC C12N5/00,C12N15/00
CC Strandedness: Single;
CC Topology: Linear;
CC Ligand receptors and utilization thereof
FH Key Location/Qualifiers
FT CDS 184..1194.
FEATURES
  source
    1..2581
      /organism="unidentified"
      /mol_type="genomic DNA"
      /db_xref="taxon:32644"
ORIGIN
Query Match 6.3%; Score 87; DB 6; Length 2581;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 2492 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 2551
QY 1322 ACGGTGATGCGAGTCATAGCTTT 1348
Db 2552 ACGGTGATGCGAGTCATAGCTTT 2578
```

```

RESULT 15
AR137825/c
LOCUS AR137825 1330 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 29 from patent US 6197561.
ACCESSION AR137825
VERSION AR137825.1 GI:14479334
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
FEATURES
  Location/Qualifiers
    1..1330
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Query Match 6.0%; Score 82; DB 6; Length 1330;
Best Local Similarity 100.0%; Pred. No. 9.3e-31;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1270 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 94 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 35
QY 1330 ATGCGAGTCATAGCTCTTTCTA 1351
Db 34 ATGCGAGTCATAGCTCTTTCTA 13
Search completed: October 2, 2004, 07:39:50
Job time : 5271 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 05:49:02 ; Search time 127 Seconds
(without alignments)
5995.218 Million cell updates/sec

Title: US-09-993-808B-1

Perfect score: 1372

Sequence: 1 cccacgcgtccgacgcaag.....agtgcacctaattcattc 1372

Scoring table: OLIGO NUC

Gapop_60.0, Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	6.9	1255	US-09-118-442-31	Sequence 31, Appl
2	95	6.9	1255	US-09-677-064-31	Sequence 31, Appl
3	95	1581	1	US-08-383-756-1	Sequence 1, Appli
4	95	6.9	1581	US-08-460-898-1	Sequence 1, Appli
5	90	6.6	944	US-09-227-357-122	Sequence 122, App
6	87	6.3	2581	US-09-013-634-1	Sequence 1, Appli
7	82	6.0	1330	US-09-118-442-29	Sequence 29, Appl
8	82	6.0	1330	US-09-677-064-29	Sequence 29, Appl
9	80	5.8	1984	US-09-904-615-46	Sequence 584, App
10	79	5.8	826	US-09-620-312D-584	Sequence 584, App
11	79	5.8	907	US-09-620-312D-585	Sequence 585, App
12	79	5.8	955	US-09-620-312D-228	Sequence 228, App
13	79	5.8	1200	US-09-620-312D-564	Sequence 564, App
14	79	5.8	2544	US-09-483-371-1	Sequence 1, Appli
15	79	5.8	2544	US-09-957-156-1	Sequence 1, Appli
16	78	5.7	1243	US-09-620-312D-1092	Sequence 1092, Ap
17	78	5.7	2638	US-09-042-785A-22	Sequence 22, Appl
18	77	5.6	1689	US-09-053-374A-4	Sequence 4, Appli
19	76	5.5	657	US-09-620-312D-543	Sequence 543, App
20	76	5.5	1067	US-09-620-312D-853	Sequence 853, App
21	76	5.5	1663	US-09-620-312D-6	Sequence 6, Appli
22	76	5.5	2085	US-09-620-312D-1002	Sequence 1002, Ap
23	76	5.5	3396	US-09-668-680-6	Sequence 6, Appli
24	76	5.5	3423	US-09-668-680-7	Sequence 7, Appli
25	75	5.5	831	US-09-904-615-25	Sequence 25, Appl
26	75	5.5	2202	US-09-396-149-3	Sequence 3, Appli
27	73	5.3	322	US-08-956-171E-1520	Sequence 1520, Ap

ALIGNMENTS

RESULT 1

US-09-118-442-31
; Sequence 31, Application US/09118442B

; Patent No. 6197561

; GENERAL INFORMATION:

; APPLICANT: Martino-Catt, Susan J.

; APPLICANT: Wang, Hongyu

; APPLICANT: Beach, Larry R.

; APPLICANT: Wang, Xun

; APPLICANT: Bowen, Benjamin A.

; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in

; FILE REFERENCE: 0706

; CURRENT APPLICATION NUMBER: US/09/118,442B

; EARLIER FILING DATE: 1998-07-17

; EARLIER APPLICATION NUMBER: 60/055,446

; EARLIER FILING DATE: 1997-08-11

; EARLIER APPLICATION NUMBER: 60/055,526

; EARLIER FILING DATE: 1997-08-08

; EARLIER APPLICATION NUMBER: 60/053,944

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 31

; LENGTH: 1255

; TYPE: DNA

; ORGANISM: Zea mays

US-09-118-442-31

Query Match 6.9%; Score 95; DB 3; Length 1255;

Best Local Similarity 100.0%; Pred. No. 7.3e-28;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATCG 1334

Db 1161 AAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATCG 1220

QY 1335 ACGTCATAGCTTCTTATAGTGTACCTAAATTCA 1369

Db 1221 ACGTCATAGCTTCTTATAGTGTACCTAAATTCA 1255

RESULT 2

US-09-677-064-31

; Sequence 31, Application US/09677064

; Patent No. 6291224

; GENERAL INFORMATION:

; APPLICANT: Martino-Catt, Susan J.

; APPLICANT: Wang, Hongyu

; APPLICANT: Beach, Larry R.

;; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
;; Plants and Uses Thereof

;; FILE REFERENCE: 0706D

;; CURRENT APPLICATION NUMBER: US/09/677,064

;; CURRENT FILING DATE: 2000-09-29

;; PRIOR APPLICATION NUMBER: 60/055,446

;; PRIOR FILING DATE: 1997-08-11

;; PRIOR APPLICATION NUMBER: 60/055,526

;; PRIOR FILING DATE: 1997-08-08

;; PRIOR APPLICATION NUMBER: 60/053,944

;; PRIOR FILING DATE: 1997-07-28

;; PRIOR APPLICATION NUMBER: 09/118,442

;; PRIOR FILING DATE: 1998-07-17

;; NUMBER OF SEQ ID NOS: 31

;; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 31

;; LENGTH: 1255

;; TYPE: DNA

;; ORGANISM: Zea mays

;; US-09-677-064-31

Query Match 6.9%; Score 95; DB 3; Length 1255;

Best Local Similarity 100.0%; Pred. No. 7.3e-28;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1275 AAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATGCG 1334

Db 1161 AAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATGCG 1220

Qy 1335 ACGTCATAGCTCTTCTATAGTGTCACCTAAATTC A 1369

Db 1221 ACGTCATAGCTCTTCTATAGTGTCACCTAAATTC A 1255

RESULT 3

US-08-383-756-1

; Sequence 1, Application US/08383756

; Patent No. 5654495

; GENERAL INFORMATION:

; APPLICANT: Dehesh, Katayoon

; APPLICANT: Voelker, Toni Alois

; APPLICANT: Hawkins, Deborah

; APPLICANT: Davies, Huw Maelor

; TITLE OF INVENTION: Production of Myristate in Plant Cells

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Calgene, Inc.

; STREET: 1920 Fifth Street

; CITY: Davis

; STATE: CA

; COUNTRY: USA

; ZIP: 95616

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh 7.0

; SOFTWARE: Microsoft Word 5.1(a)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/383,756

; FILING DATE: 02-FEB-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/261,695

; FILING DATE: 16-JUN-94

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/10814

; FILING DATE: 29-OCT-93

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 07/968,971

; FILING DATE: 30-OCT-92

; ATTORNEY/AGENT INFORMATION:

; NAME: Elizabeth Lassen

; REGISTRATION NUMBER: 31,845

; NAME: Donna E. Scherer

; REGISTRATION NUMBER: 34,719

; NAME: Carl J. Schwedler

; REGISTRATION NUMBER: 36,924

; REFERENCE/DOCKET NUMBER: CGNE 111

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (916) 753-6313

; TELEFAX: (916) 753-1510

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1581 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; US-08-383-756-1

Query Match 6.9%; Score 95; DB 1; Length 1581;

Best Local Similarity 100.0%; Pred. No. 7e-28;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1275 AAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATGCG 1334

Db 1479 AAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATGCG 1538

Qy 1335 ACGTCATAGCTCTTCTATAGTGTCACCTAAATTC A 1369

Db 1539 ACGTCATAGCTCTTCTATAGTGTCACCTAAATTC A 1573

RESULT 4

US-08-460-898-1

; Sequence 1, Application US/08460898

; Patent No. 5850022

; GENERAL INFORMATION:

; APPLICANT: Dehesh, Katayoon

; APPLICANT: Voelker, Toni Alois

; APPLICANT: Hawkins, Deborah

; APPLICANT: Davies, Huw Maelor

; TITLE OF INVENTION: Production of Myristate in Plant Cells

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Calgene, Inc.

; STREET: 1920 Fifth Street

; CITY: Davis

; STATE: CA

; COUNTRY: USA

; ZIP: 95616

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh 7.0

; SOFTWARE: Microsoft Word 5.1(a)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/460,898

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/383,756

; FILING DATE: 02-FEB-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/261,695

; FILING DATE: 16-JUN-94

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/10814

; FILING DATE: 29-OCT-93

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 07/968,971

```

; FILING DATE: 30-OCT-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 111-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-08-460-898-1

Query Match      6.9%; Score 95; DB 2; Length 1581;
Best Local Similarity 100.0%; Pred.No. 7e-28;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1275  AAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCATGCG 1334
Db      1479  AAAAAAAAAAAAAAGGCGCGCCCTCTAGAGGATCCAGCTTACGTACGGTGCATGCG 1538

QY      1335  ACGTCATAGCTCTTCTATAGTGTCACCTAAATTC 1369
Db      1539  ACGTCATAGCTCTTCTATAGTGTCACCTAAATTC 1573

```

RESULT 5

US-09-227-357-122

Sequence 122, Application US/09227357

Patent No. 6342581

GENERAL INFORMATION:

APPLICANT: Fischer et al.

TITLE OF INVENTION: 123 Human Secreted Proteins

FILE REFERENCE: P2010P1

CURRENT APPLICATION NUMBER: US/09/227,357

CURRENT FILING DATE: 1999-01-08

EARLIER APPLICATION NUMBER: PCT/US98/13684

EARLIER FILING DATE: 1998-07-07

EARLIER APPLICATION NUMBER: 60/051,926

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,793

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,925

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,929

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,803

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,732

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,931

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,932

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,916

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,930

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,918

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,920

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,733

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,795

```

; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (932)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (942)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (944)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-227-357-122

Query Match      6.6%; Score 90; DB 4
Best Local Similarity 100.0%; Pred. No. 6.5e
Matches 90; Conservative 0; Mismatches

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGC
Db 835 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGC

QY 1322 ACGGTGATCGACGCTCATAGCTCTTCTA 1351
Db 895 ACGGTGATCGACGCTCATAGCTCTTCTA 924

RESULT 6
US-09-013-634-1
Sequence 1, Application US/09013634

```

```
; Patent No. 5945307
; GENERAL INFORMATION:
; APPLICANT: M. Alexandra Glucksmann and Keith Robison
; TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,634
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Jean M. Silveri
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MMI-036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 184..1194
; US-09-013-634-1

Query Match 6.3%; Score 87; DB 2; Length 2581;
Best Local Similarity 100.0%; Pred. No. 7.7e-25;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 2492 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 2551

QY 1322 ACGGTGATCGGAGCTCATGCTTT 1348
Db 2552 ACGGTGATCGGAGCTCATGCTTT 2578

RESULT 7
US-09-118-442-29/c
; Sequence 29, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11

; Patent No. 5945307
; GENERAL INFORMATION:
; APPLICANT: M. Alexandra Glucksmann and Keith Robison
; TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,634
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Jean M. Silveri
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MMI-036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 184..1194
; US-09-013-634-1

Query Match 6.0%; Score 82; DB 3; Length 1330;
Best Local Similarity 100.0%; Pred. No. 7.3e-23;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 94 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 35

QY 1330 ATGCGAGCTCATGCTTCTTA 1351
Db 34 ATGCGAGCTCATGCTTCTTA 13

RESULT 8
US-09-677-064-29/c
; Sequence 29, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 1330
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-677-064-29

Query Match 6.0%; Score 82; DB 3; Length 1330;
Best Local Similarity 100.0%; Pred. No. 7.3e-23;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 94 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 35

QY 1330 ATGCGAGCTCATGCTTCTTA 1351
Db 34 ATGCGAGCTCATGCTTCTTA 13

RESULT 9
US-09-904-615-46
; Sequence 46, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```
/ TITLE OF INVENTION: 49 Human Secreted Proteins
/ FILE REFERENCE: P2032P1
/ CURRENT APPLICATION NUMBER: US/09/904,615
/ CURRENT FILING DATE: 2001-07-16
/ PRIOR APPLICATION NUMBER: 09/511,554
/ PRIOR FILING DATE: 2000-02-23
/ PRIOR APPLICATION NUMBER: 60/097,917
/ PRIOR FILING DATE: 1998-08-25
/ PRIOR APPLICATION NUMBER: 60/098,634
/ PRIOR FILING DATE: 1998-08-31
/ NUMBER OF SEQ ID NOS: 170
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 46
/ LENGTH: 1984
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (106)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-46

Query Match      5.8%; Score 80; DB 4; Length 1984;
Best Local Similarity 100.0%; Pred. No. 46-22;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATGCG 1334
Db 1898 AAAAAAAAAAAGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATGCG 1957

QY 1335 ACGTCATAGCTCTCTATAG 1354
Db 1958 ACGTCATAGCTCTCTATAG 1977

RESULT 10
US-09-620-312D-584
/ Sequence 584, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yunding
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 584
/ LENGTH: 907
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (225)..(695)
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 584
/ LENGTH: 826
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
```

```
/ NAME/KEY: CDS
/ LOCATION: (225)..(614)
US-09-620-312D-584

Query Match      5.8%; Score 79; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATGCGATAGCTCTTCT 1350
Db 743 GCGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATGCGATAGCTCTTCT 802

QY 1351 ATAGTGTCACTAAATTC A 1369
Db 803 ATAGTGTCACTAAATTC A 821

RESULT 11
US-09-620-312D-585
/ Sequence 585, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yunding
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 585
/ LENGTH: 907
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (225)..(695)
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 585
/ LENGTH: 907
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (225)..(695)
US-09-620-312D-585

Query Match      5.8%; Score 79; DB 4; Length 907;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATGCGATAGCTCTTCT 1350
Db 824 GCGGCGCGCTCTAGAGATCCAAAGCTTACGTACGCGTGCATGCGATAGCTCTTCT 883

QY 1351 ATAGTGTCACTAAATTC A 1369
Db 884 ATAGTGTCACTAAATTC A 902

RESULT 12
```

```
US-09-620-312D-228/c
; Sequence 228, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 228
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)...(583)
US-09-620-312D-228

Query Match 5.8%; Score 79; DB 4; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1291 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 1350
Db 120 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 61

Qy 1351 ATAGTGTCACTAAATTTCA 1369
Db 60 ATAGTGTCACTAAATTTCA 42

RESULT 13
US-09-620-312D-564/c
; Sequence 564, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 228
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)...(583)
US-09-620-312D-228

Query Match 5.8%; Score 79; DB 4; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1291 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 1350
Db 120 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 61

Qy 1351 ATAGTGTCACTAAATTTCA 1369
Db 60 ATAGTGTCACTAAATTTCA 42

RESULT 13
US-09-620-312D-564/c
; Sequence 564, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 564
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (559)...(957)
US-09-620-312D-564

Query Match 5.8%; Score 79; DB 4; Length 1200;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1291 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 1350
Db 86 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 27

Qy 1351 ATAGTGTCACTAAATTTCA 1369
Db 26 ATAGTGTCACTAAATTTCA 8

RESULT 14
US-09-483-371-1
; Sequence 1, Application US/09483371
; Patent No. 6309869
; GENERAL INFORMATION:
; APPLICANT: Debbie S. Yaver
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; TITLE OF INVENTION: Polypeptides Having Acid Phosphatase
; TITLE OF INVENTION: Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5791.200-US
; CURRENT APPLICATION NUMBER: US/09/483,371
; CURRENT FILING DATE: 2000-01-14
; EARLIER APPLICATION NUMBER: 09/231,612
; EARLIER FILING DATE: 1999-01-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2544
; TYPE: DNA
; ORGANISM: Fusarium
US-09-483-371-1

Query Match 5.8%; Score 79; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 9.2e-22;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1291 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 1350
Db 2305 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACCGGTGCATGCGACGTCATAGCTCTTCT 2364

Qy 1351 ATAGTGTCACTAAATTTCA 1369
Db 2365 ATAGTGTCACTAAATTTCA 2383

RESULT 15
```

```

US-09-957-156-1
; Sequence 1, Application US/09957156
; Patent No. 6667169
; GENERAL INFORMATION:
; APPLICANT: Debbie S. Yaver
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; TITLE OF INVENTION: Polypeptides Having Acid Phosphatase
; TITLE OF INVENTION: Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5791.200-US
; CURRENT APPLICATION NUMBER: US/09/957,156
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/483,371
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2544
; TYPE: DNA
; ORGANISM: Fusarium
US-09-957-156-1

```

```

Query Match          5.8%; Score 79; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 9.2e-22;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGCGCGCTTAGAGGATCCAGCTTAGCTAGCGCGTGCATGCGAGCTCATAGCTCTTCT 1350
Db 2305 GCGGCGCGCTTAGAGGATCCAGCTTAGCTAGCGCGTGCATGCGAGCTCATAGCTCTTCT 1350

QY 1351 ATAGTGTACCTAAATTCA 1369
Db 2365 ATAGTGTACCTAAATTCA 2383

```

Search completed: October 2, 2004, 08:37:50
Job time : 128 secs

This Page Blank (uspto)


```
QY 421 GGCTGCTGATGGCGCTGCAGACAGCAGGGCGGGCGCTCGCGCGGGCTCTCGCGTTG 480
Db 421 GGCTGCTGATGGCGCTGCAGACAGCAGGGCGGGCGCTCGCGCGGGCTCTCGCGTTG 480
QY 481 CTCACGACGGCGCTGCTCGTGAACCTTGGGCTTGGGGGTCAGCGGGGAGCCACACTG 540
Db 481 CTCACGACGGCGCTGCTCGTGAACCTTGGGCTTGGGGGTCAGCGGGGAGCCACACTG 540
QY 541 CCCTCTCTACGACGCTGCAGAGCTGCGCGGGATCAGTCTCTGTTGATGCTCGCGGC 600
Db 541 CCCTCTCTACGACGCTGCAGAGCTGCGCGGGATCAGTCTCTGTTGATGCTCGCGGC 600
QY 601 GAGCAACTCGGAGCGGCCAGACCGCGAGAGCGGAGAGACGACGACGACGCGCGGC 660
Db 601 GAGCAACTCGGAGCGGCCAGACCGCGAGAGCGGAGAGACGACGACGACGCGCGGC 660
QY 661 GCACGGGAGCTCAGCGATCTGAGTCGATCTGGCGGGGACAAAGACTGGCCGCTGCT 720
Db 661 GCACGGGAGCTCAGCGATCTGAGTCGATCTGGCGGGGACAAAGACTGGCCGCTGCT 720
QY 721 ACCGGCGGCAACCGCGCTCGGAGCTGATCGTCCGCGGACACAGAGATCCAGAGTT 780
Db 721 ACCGGCGGCAACCGCGCTCGGAGCTGATCGTCCGCGGACACAGAGATCCAGAGTT 780
QY 781 CTTTCGCGCGCGCGAGCGCGGCCAGGCCCAAGCGCTTTGCTTCCAACTTCAACTT 840
Db 781 CTTTCGCGCGCGCGAGCGCGGCCAGGCCCAAGCGCTTTGCTTCCAACTTCAACTT 840
QY 841 GCTTCGCGCGCGCTGCTCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 900
Db 841 GCTTCGCGCGCGCTGCTCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 900
QY 901 CTTGAAGCGAGCTGCTCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 960
Db 901 CTTGAAGCGAGCTGCTCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 960
QY 961 CCCCCCCCCCAACAAATACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 CCCCCCCCCCAACAAATACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 GCTTAAGCTAACCAACCACTTCTGTCGAAATGATGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 GCTTAAGCTAACCAACCACTTCTGTCGAAATGATGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 TAGCAGAGCTAGTTATTTATTTAGTACTTATTTAGTACTTATTTAGTACTTAT 1140
Db 1081 TAGCAGAGCTAGTTATTTATTTAGTACTTATTTAGTACTTATTTAGTACTTAT 1140
QY 1141 GATAGTCAATCGGCTTACTGTGTAATTTCTCTCATGCTATCTTTAGATGGATTTAATC 1200
Db 1141 GATAGTCAATCGGCTTACTGTGTAATTTCTCTCATGCTATCTTTAGATGGATTTAATC 1200
QY 1201 GTCTTAATTTATTTAGTACTGAGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 GTCTTAATTTATTTAGTACTGAGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 CAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 CAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 TAGCGGTGATGCGAGCTCATAGCTTCTTATAGTGTACCTAAATTCATTC 1372
Db 1321 TAGCGGTGATGCGAGCTCATAGCTTCTTATAGTGTACCTAAATTCATTC 1372
```

RESULT 3

```
US-10-425-114-18379/c
; Sequence 18379, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18379
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI
US-10-425-114-18379
```

Query Match 26.2%; Score 360; DB 13; Length 985;

Best Local Similarity 99.8%; Pred. No. 2.7e-164;

Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 549 ACACGCTGAGAGGCTGCGGGATCACGCTCTGTGTGATGTCTCGCGCGGAGCAACT 608
Db 674 ACACGCTGAGAGGCTGCGGGATCACGCTCTGTGTGATGTCTCGCGCGGAGCAACT 615
QY 609 CCGGAGCGCGCCAGACCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668
Db 614 CCGGAGCGCGCCAGACCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 655
QY 669 AGCTACGCGATCTGAGTCTGATCTGCGGGGACAAAGACTGGCCCGTCTGCTACCGCGG 728
Db 554 AGCTACGCGATCTGAGTCTGATCTGCGGGGACAAAGACTGGCCCGTCTGCTACCGCGG 495
QY 729 CAACGCGGCTGCGGAGCTGATGTCGCCGCCAGCACACGAGATCCAGAGATCTTTCGCGG 788
Db 494 CAACGCGGCTGCGGAGCTGATGTCGCCGCCAGCACACGAGATCCAGAGATCTTTCGCGG 435
QY 789 CCGCGGAGCGCGCCAGCGCCAGCGCTTTGCTTCCAACTTCAACTTCAACTTCAACTT 848
Db 434 CCGCGGAGCGCGCCAGCGCCAGCGCTTTGCTTCCAACTTCAACTTCAACTTCAACTT 375
QY 849 GCGTGCCTCTCAGCGCGCGCGCGCTTCTGAGTGGCGCGCGGTGCTCAGCATCTGAAGCG 908
Db 374 GCGTGCCTCTCAGCGCGCGCGCGCTTCTGAGTGGCGCGCGGTGCTCAGCATCTGAAGCG 315
QY 909 AGCGTGCCTCTCAGCGCGCGCGCGCTTCTGAGTGGCGCGCGGTGCTCAGCATCTGAAGCG 959
Db 314 AGCGTGCCTCTCAGCGCGCGCGCGCTTCTGAGTGGCGCGCGGTGCTCAGCATCTGAAGCG 264
```

RESULT 4

```
US-10-425-114-34947
; Sequence 34947, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34947
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017355P05_FLI
```



```

; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: misc_feature
; LOCATION: (1347)..(1347)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: misc_feature
; LOCATION: (1527)..(1527)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: misc_feature
; LOCATION: (1533)..(1533)
; OTHER INFORMATION: n equals a,t,g, or c
;
US-10-023-896-42

```

	Query Match	7.9%	Score 108;	DB 15;	Length 1557;
	Best Local Similarity	100.0%;	Pred. No. 5e-42;		
	Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1262	AAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGT	1321		
Db	1211	AAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGT	1270		
QY	1322	ACCGGTGCATGCGACGTCATAGCTCTCTATAGTGTCAACCTAAATTC	1369		
Db	1271	ACCGGTGCATGCGACGTCATAGCTCTCTATAGTGTCAACCTAAATTC	1318		

```

RESULT 8
US-09-925-297-261
: Sequence 261, Application US/09925297
: Patent No. US20020081859A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: P105

```

	Query Match	7.9%	Score 108;	DB 9;	Length 2116;
	Best Local Similarity	100.0%;	Pred. No. 4.8e-42;		
	Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1262	AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCGCTCTAGAGATCCAAGCTTACGT	1321		
db	1977	AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCGCTCTAGAGATCCAAGCTTACGT	2036		

Qy	1322	ACGGTGCATGCGAGTCATAGCTCTTCTATAGTGTCACTAAATTCA	1369
		ACGGTGCATGCGAGTCATAGCTCTTCTATAGTGTCACTAAATTCA	
Db	2037	ACGGTGCATGCGAGTCATAGCTCTTCTATAGTGTCACTAAATTCA	2084
RESULT 9			
US-10-023-896-31			
; Sequence 31, Application US/10023896			
; Publication No. US20030027776A1			

	Query Match	7.9%	Score 108;	DB 15;	Length 2116;
	Best Local Similarity	100.0%;	Pred. No. 4.8e-42;		
	Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1262	AAAAAAAAAAAAAAAAAAAAAAGGGCGGCGCTCTAGAGGATCCAAAGCTTACGT	1321		
Db	1977	AAAAAAAAAAAAAAAAAAAAAAGGGCGGCGCTCTAGAGGATCCAAAGCTTACGT	2036		
QY	1322	ACGGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA	1369		
Db	2037	ACGGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA	2084		

```

RESULT 10
US/10-106-698-519
; Sequence 519, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer
; FILE REFERENCE: PA00591
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137

```

PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 519
LENGTH: 701
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (9)..(9)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (34)..(34)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (600)..(600)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (637)..(637)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (647)..(647)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (676)..(676)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (691)..(691)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (700)..(700)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-519

Query Match 7.7%; Score 106; DB 15; Length 701;
Best Local Similarity 100.0%; Pred. No. 5.1e-41;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1264 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTAC 1323
Db 459 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTAC 518
Qy 1324 GCGTCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTCA 1369
Db 519 GCGTCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTCA 564

RESULT 11
US-10-264-049-133/c
Sequence 133, Application US/10264049
Publication No. US2004000579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 133
LENGTH: 845
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
NAME/KEY: misc feature
LOCATION: (691)..(691)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (845)..(845)
OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-133

Query Match 7.5%; Score 103; DB 16; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.4e-39;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1267 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCG 1326
Db 178 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCG 119
Qy 1327 TGCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTCA 1369
Db 118 TGCATGCGACGTCATAGCTCTTCTATAGTGCACCTAAATTCA 76

RESULT 12
US-10-264-237-1380/c
Sequence 1380, Application US/10264237
Publication No. US20040009491A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P131P1
CURRENT APPLICATION NUMBER: US/10/264,237
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/16450
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,515
PRIOR FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 2876
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 1380
LENGTH: 2049
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (12)..(12)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (48)..(48)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (322)..(322)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (1819)..(1819)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (1987)..(1987)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (2021)..(2021)
OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1380

Query Match 7.2%; Score 99; DB 16; Length 2049;
Best Local Similarity 100.0%; Pred. No. 1.1e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1271 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCAT 1330
Db 321 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCAT 262
QY 1331 TGGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA 1369
Db 261 TGGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA 223

RESULT 13

US-10-050-704-66
; Sequence 66, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-050-704-66

Query Match 7.1%; Score 98; DB 15; Length 2657;

Best Local Similarity 100.0%; Pred. No. 3.4e-37;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1272 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCAT 1331
Db 2307 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCAT 2366
QY 1332 GCGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA 1369
Db 2367 GCGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA 2404

RESULT 14

US-10-798-512-66
; Sequence 66, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; PRIOR FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66

; LENGTH: 2657

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (179)

; OTHER INFORMATION: n equals a,t,g, or c

US-10-798-512-66

Query Match

Best Local Similarity 7.1%; Score 98; DB 17; Length 2657;

Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1272 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCAT 1331

Db 2307 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCAT 2366

QY 1332 GCGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA 1369

Db 2367 GCGACGTCATAGCTCTTCTATAGTGTCACCTAAATTCA 2404

RESULT 15

US-10-106-698-2395
; Sequence 2395, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2395
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (505)..(505)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (519)..(519)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (533)..(533)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2395

Query Match

Best Local Similarity 7.1%; Score 97; DB 15; Length 544;

Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1273 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCATG 1332

Db 236 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGTGCATG 295

Mon Oct 4 08:49:08 2004

QY 1333 CGAGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 296 CGAGTCATAGCTCTTCTATAGTGTACCTAAATTCA 332

Search completed: October 2, 2004, 10:11:41
Job time : 704 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	178	13.7	196	2	T09968	cyclin-dependent k
2	149	11.4	327	2	T00797	hypothetical prote
3	129.5	9.9	1156	2	T34852	probable secreted
4	128	9.8	801	2	T29018	hypothetical prote
5	124	9.5	825	1	EDBXD	immediate-early pr
6	121.5	9.3	195	2	H96532	hypothetical prote
7	121.5	9.3	1151	2	T18535	high molecular mas
8	121.5	9.3	2639	2	T31328	fibronin - Chinese
9	121	9.3	191	2	T01132	cyclin-dependent k
10	115	8.8	1414	1	S23809	collagen alpha 2(I
11	113.5	8.7	467	2	E70976	hypothetical prote
12	113	8.7	1541	2	T02831	AAA protein L4171.
13	113	8.7	4391	2	A38096	perlecan precursor
14	112	8.6	1446	1	A45344	immediate-early pr
15	110.5	8.5	1077	2	A4067	serine-rich protei
16	109.5	8.4	294	2	S13141	hypothetical prote
17	107.5	8.2	490	2	T09084	phosphatidylinosit
18	107.5	8.2	698	2	T01209	starch synthase (E
19	107	8.2	278	2	S44796	F09G8.6 protein -
20	107	8.2	572	2	T08509	trbL protein - Ent
21	107	8.2	632	2	T00084	hypothetical prote
22	106.5	8.2	456	2	G97677	hypothetical prote
23	106	8.1	306	2	F84276	forminoglutamate
24	106	8.1	738	2	E87627	hypothetical prote
25	106	8.1	1147	2	T35781	hypothetical prote
26	105.5	8.1	1122	2	G64887	probable tail fibe
27	105	8.1	237	2	A88640	protein C34H4.4 [i
28	105	8.1	583	1	S22544	transcription fact
29	104.5	8.0	413	2	AH2743	conserved hydrophe


```

Best Local Similarity   24.3%; Pred. No. 0.13;
Matches    42; Conservative    27; Mismatches    53; Indels    51; Gaps    7

QY  114 SRCSTASSVNLGLGGQGSHTCRSYDAAEAGG-----DHVLVDVSAASNSGSG 162
Db   33 TRLUSSSSSSL-----AYSVDGGGFCVSALSEEEDHLSSISSSGCSSET 80

QY  163 PDRRRRTTPSSRAHGELSDLE-SDL-----AGHKTGPSLPAATPADEL----- 205
Db   81 NEIATRLPFSDLAH-EISETEISTLTNNFKOGISSENLGETAEMDSATTMRDQRK 139

QY  206 -----IVPPAHEIOEFFAAEAQAOKPASKYNFDFVGVPFLDDAGGREWF 250
Db   140 TEKKKMEKSPTQAELODFTSAAERYEQKRFTKYNYDIVNDTPLE--GRYQW 190

RESULT 7
TI8535
high molecular mass nuclear antigen - chicken (fragment)
C/Species: Gallus gallus (chicken)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Acession: TI8535
R/Shimada, K.; Harata, M.; Mizuno, S.
J. Cell Sci. 110, 3031-3041, 1997
A/Title: A nuclear matrix-associated high molecular mass nuclear antigen, HMNA,
A/Reference number: Z18955; MUID:9803440; PMID:9365273
A/Acession: TI8535
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-1151 <SHI>
A/Cross-references: EMBL:D88440; NID:d1177138; PID:d1025045; PIDN:BAA24137.1

Query Match          9.3%; Score 121.5; DB 2; Length 1151;
Best Local Similarity 22.5%; Pred. No. 0.77;
Matches    61; Conservative    24; Mismatches   105; Indels    81; Gaps    7;

QY  16 VAAVEVTQ-----VVGVRTSRGAATAATGGVAKVRKRKPAGEPAAAVSAGG 63
Db   861 MAATNTQTTPPMMAASTPQSITPMGAATTQSPMGCAATTQSPFMGCASITPQAPTTVAGSP-- 918

QY  64 DGSGCYIHLSRMFLFWAPPQPQP-----SVDSVPTPVAAADGAAGQGCGAALAA 111
Db   919 -----TTPPPPIPPSPTAQTSPQPMKSPPPPDPKAPSAAAQTSPAHHVA 962

QY  112 -----GLSRCSTASSVNLGLGGQGSHTCRSYDAAEAGGDHVLVDVSAAS 157
Db   963 NASPGVTAVSPAPIGVVTEASPADGARLSPGTAAITDGPKASPAATADVTEAATDVTAATA 1022

QY  158 -----NSGSGPDRERRTTPSSRAHGELSDLESIDLGHKTGPSLFA----- 198
Db   1023 TAVPAEAAPTKAKRSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1082

QY  199 ----ATPAELIVPPAHE--IQEFFAAEEAA 223
Db   1083 GOOQMTPGAQSVPPVTEAAVQEAAAAAAAAA 1113

RESULT 8
T31328
fibroin - Chinese oak silkmoth
C/Species: Antheraea pernyi (Chinese oak silkmoth)
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C/Acession: T31328
R/Sezutsu, H.; Tamura, T.; Yukuhiro, K.
submitted to the EMBL Data Library, August 1998
A/Description: Characterization of the full length fibroin gene of a wild silkworm
A/Reference number: Z20995
A/Acession: T31328
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2639 <SEZ>
A/Cross-references: EMBL:AF083334; NID:g3450882; PID:g3450883; PIDN:AAC32606.1
C/Genetics:

```


F;3845-3880/Domain: EGF homology <EGF1>
F;3888-3921/Domain: EGF homology <EGF>
F;3953-4106/Domain: laminin G repeat homology <LG2>
F;4147-4175/Domain: EGF homology <EGF2>
F;4149-4151/Region: motor neuron attachment (L-R-E) motif
F;4299-4301/Region: motor neuron attachment (L-R-E) motif
F;65,71,76/Binding site: heparan sulfate (Ser) (covalent) #status predicted
F;89,554,1755,2121,3072,3105,3279,3780,3836,4068/Binding site: carbohydrate (Asn) (coval)
F;2395,3933,4179/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 8.7%; Score 113; DB 2; Length 4391;
Best Local Similarity 26.5%; Pred. No. 12;
Matches 75; Conservative 27; Mismatches 83; Indels 98; Gaps 16;

Qy 2 GKVMKCRGAAGAEVAEVNT-----QVVG-VRTSRGAATGG-----VA 41
Db 2499 GEYCRVVGSGTQASVLTITQRLSGSHGQGVAYPVRISSASLANGLTDLNCLVA 2558
Qy 42 KVAPR-----RKEAP-----AGEPAAVVSAG----- 62
Db 2559 SQPHTITWYKRGSLPSRHOIVGSLRIPOVTPADSGEYVCHVSNAGSRETSLIVTIQ 2618
Qy 63 GDGSGCYHLRSMFMAPPQPSVDSVPTPVEA-----ADGAAGCQGAALAGLSRCS 118
Db 2619 GSGSS---HVPS-----VSPPIRIES--SSPTWEGQTLDLNCVVARQPQAITWYKRG 2669
Qy 119 TAS-----SVNLGLG---GQSGHTCR---SYDAAEAGGDHVLVDVSAANSNGSGGPDRE 166
Db 2670 LPSRHQTHGSHLRHLMQSVASGEVVCRRANNIDALEAS---IVISVPSAGSPSAFOSS 2726
Qy 167 ---RRETPSSRAHGELSDLESLAG-----HKTGPSLPA 198
Db 2727 MPRIESSSHVAEGETDLNCVPGQAHQVTHWKRGSLPS 2769

RESULT 14

A45344
Immediate-early protein - suid herpesvirus 1 (strain Kaplan)
C:Species: suid herpesvirus 1
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C:Accession: A45344
R:Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzler, M.
Virology 179, 365-377, 1990
A>Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented op
A:Reference number: A45344; MUID:91021039; PMID:2171211
A:Accession: A45344
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1446 <VLC>
A:Cross-references: GB:M34651; NID:g334070; PIDN:AAA47470.1; PID:g334071
C:Superfamily: herpesvirus immediate-early protein IE175
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 8.6%; Score 112; DB 1; Length 1446;
Best Local Similarity 26.8%; Pred. No. 4.4;
Matches 64; Conservative 19; Mismatches 78; Indels 78; Gaps 11;

Qy 9 RCAGAEVAAEVNT-----QVGVVTRSR-----AAATGGVAKVAPRRKRAPAGEPAAA 58
Db 745 RQQAADSVLVARTVAPLVVYSDGARAEAAWTVAALFAPANVAALAEAAAPGPA 804
Qy 59 VSAGDGGSCYHLRSMFMAPPQPSVDSVPTPVEAADGA-AGCQGAALAGLSRCS 117
Db 805 EPAPG-----LPPLWPEQGLVVPAPAPAAAGAPSLPGSGSPSPASTKS 849
Qy 118 S-----TASVNLGLGQSGHTCRSYDAAEAGGDHVLVDVSAANSNGSGPD-----RERR 168
Db 850 SSSSTKSSSTKSGSGSG-----YASSPAAGPDPAPEPRKKK 887
Qy 169 ETPSSRAHGELSD-----LESGLAGK-----TGP-----SI-----PAATPAELI 206
Db 888 RRAPGARPEGDEEGLSGAALRGCHGRDDEDRGRPRKRSLSGLGPAPDPAPALL 946

RESULT 15

A44067
serine-rich protein hairless - fruit fly (Drosophila melanogaster)
N:Alternate names: 109K basic protein H
C:Species: Drosophila melanogaster
C>Date: 10-Jun-1993 #sequence_revision 26-Feb-1999 #text_change 26-Feb-1999
C:Accession: A44067; A58929; S33412; S24639
R:Bang, A.G.; Posakony, J.W.
Genes Dev. 6, 1752-1769, 1992
A>Title: The Drosophila gene Hairless encodes a novel basic protein that controls alterna
A:Reference number: A44067; MUID:92387549; PMID:1516831
A:Accession: A44067
A:Molecule type: DNA
A:Residues: 19-1077 <BAN>
A:Cross-references: GB:M95192; NID:g157621; PID:g157622
A>Note: sequence extracted from NCBI backbone (NCBIN:112622, NCBI:P:112623)
R:Preiss, A.
submitted to the EMBL Data Library, May 1994
A:Description: Hairless, a Drosophila gene involved in neural development, encodes a nove
A:Reference number: A58929
A:Accession: A58929
A:Molecule type: mRNA
A:Residues: 1-1077 <PRE>
A:Cross-references: EMBL:X67239; GB:S49642; NID:g578331; PID:g578332
R:Maier, D.; Stumm, G.; Kuhn, K.; Preiss, A.
Mech. Dev. 38, 143-156, 1992
A>Title: Hairless, a Drosophila gene involved in neural development, encodes a novel, ser
A:Reference number: S33412; MUID:93041287; PMID:1419850
A:Accession: S33412
A:Molecule type: mRNA
A:Residues: 1-150, 'A', 152-701, 'LL', 704-890, 'R', 892-963, 'RLLP', 968-973, 975-1077 <MAI>
A:Cross-references: EMBL:X67239
C:Genetics:
A:Gene: FlyBase:H: hairless
A:Cross-references: FlyBase:FBgn0001169

Query Match 8.5%; Score 110.5; DB 2; Length 1077;
Best Local Similarity 20.9%; Pred. No. 4.1;
Matches 58; Conservative 33; Mismatches 87; Indels 99; Gaps 11;
Qy 33 SAAATGGVAK-----VAPRRK-----RAPAGEPAAVVSAGDGGSCYHLRSLR 75
Db 557 SAAAGRLVVEYHTOHVSPRKRILREFKVLSDNCGVNGSGGASGGAGGK-----RSR 611
Qy 76 M-----LFMAPPPQPSVDSVPTPVEAADGAAGCQGAALAGLSRCSSTAS 121
Db 612 AKGISTSSPAGKASPMNLAPQKPS-----PSP-----GSSSSSTSPATLSTQPTRLNSSYS 664
Qy 122 SVNLGLGQSGHTCRSYDAAEAGGDHVLVDVS----- 154
Db 665 IHSLL-LGGSSGSGSSSSSSSGKCGDHPAAIISNVHFQHSNMYQPSSSSYPRALLTSPKS 723
Qy 155 ---AASNSGSGPDRERET-----TPSSRAHGELSD-----LESGLA 188
Db 724 PDVSGNSGSGKSPSHTKRSPYSGSPVDYHGFYRDYAGAGRPSISGSASQDLS 783
Qy 189 GHKTGPSLPAATPAELIVPPAHEIQEFAAAEAQA 225
Db 784 PRRSPAPATTP--RTVPKKTASIRREFASPSASS 818

Search completed: October 1, 2004, 16:54:33
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:45:15 ; Search time 24. Seconds
(without alignments)

555.415 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 1304

Sequence: 1 MGKYMKRCGAGAEVAABE.....VRGVPLDAGGFEWAPVWSI 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	9.5	825	1	ICP0 HSV2H
2	113	8.7	4391	1	PGBM HUMAN
3	112	8.6	1446	1	IE18 PRVKA
4	111.5	8.6	1115	1	TBC2 CHURE
5	110.5	8.5	1077	1	HLES DROME
6	107	8.2	278	1	VLSE_CABEL
7	106	8.1	306	1	HUTG HALN1
8	106	8.1	677	1	OGFR HUMAN
9	106	8.1	1461	1	IE18 PRVIF
10	105.5	8.1	1120	1	STFR ECOLI
11	105	8.1	553	1	FXC1 MOUSE
12	104.5	8.0	1487	1	ICP4 HSVEB
13	104.5	8.0	1487	1	ICP4 HSVBK
14	104	8.0	219	1	BASP RAT
15	104	8.0	888	1	KLTK MOUSE
16	103.5	7.9	388	1	HXAD HUMAN
17	103.5	7.9	391	1	SOX1 MOUSE
18	103.5	7.9	518	1	TPM4 DROME
19	102.5	7.9	386	1	HXAD MOUSE
20	102.5	7.9	553	1	FXC1 HUMAN
21	102.5	7.9	1174	1	CIKE DROME
22	102	7.8	502	1	SYFA HALN1
23	101.5	7.8	447	1	ODP2 RHIME
24	101.5	7.8	704	1	CL16_HUMAN
25	101	7.7	774	1	STF_LAMB
26	100.5	7.7	864	1	KLTK HUMAN
27	100.5	7.7	1093	1	AF17 HUMAN
28	100.5	7.7	1163	1	Y222 HUMAN
29	100	7.7	907	1	AI80 HUMAN
30	100	7.7	1321	1	IRS2 MOUSE
31	99.5	7.6	328	1	HXD1 HUMAN
32	99.5	7.6	775	1	ICP0 HSV11
33	99.5	7.6	2333	1	PGCA_CANFA

34	99	7.6	676	1	ICP0 HSVBJ
35	99	7.6	1186	1	HCN4_MOUSE
36	99	7.6	3707	1	PGBM_MOUSE
37	98.5	7.6	901	1	AI80_MOUSE
38	98.5	7.6	910	1	PERT BORPE
39	98	7.5	440	1	DCO DROME
40	98	7.5	1324	1	IRS2 HUMAN
41	97.5	7.5	676	1	ICP0 HSVBK
42	97.5	7.5	1733	1	VNUA_PRVKA
43	97	7.4	2038	1	FSH DROME
44	96.5	7.4	336	1	TTC9 HUMAN
45	96.5	7.4	421	1	CCG8_RAT

P29128	bovine herp
O70507	mus musculus
Q05793	mus musculus
Q61548	mus musculus
P14283	bordetella
O76324	drosophila
O9y4h2	homo sapien
P29836	bovine herp
P33485	pseudorabie
P13709	drosophila
Q92623	homo sapien
Q8vhw5	rattus norv

ALIGNMENTS

RESULT 1

ICP0 HSV2H
ID ICP0 HSV2H STANDARD; PRT; 825 AA.
AC P28284;1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICP0 (VMW118 protein).
GN RL2
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2."
RL J. Gen. Virol. 72:3057-3075(1991).
RN [2]
RP SEQUENCE FROM N.A.

RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Contains 1 RING-type zinc finger.

CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ICP0 PROTEIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D10471; BAA23427.1; -

DR EMBL; Z86099; CAB06760.1; -

DR PIR; JQ1501; EDBEXD.

DR HSP; P28990; ICHC.

DR InterPro; IPR001841; Znf_ring.

DR Pfam; PF00097; zf-C3HC4; 1.

DR SMART; SM00184; RING; 1.

DR PROSITE; PS00518; ZF_RING_1; 1.

DR PROSITE; PS00089; ZF_RING_2; 1.

DR Transcription regulation; Trans-acting factor; Activator; Zinc-finger;

DR DNA-binding.

FT DOMAIN 120 123 POLY-GLY.

FT ZN FING 126 167 RING-TYPE.

FT DOMAIN 266 271 POLY-GLY.

FT DOMAIN 292 295 POLY-SER.

FT DOMAIN 342 345 POLY-ALA.

FT DOMAIN 386 389 POLY-SER.

FT DOMAIN 395 400 POLY-GLY.

FT DOMAIN 425 428 POLY-ALA.

FT DOMAIN 590 627 POLY-SER.

```
SQ SEQUENCE 825 AA; 81986 MW; 5CEB15858553A274 CRC64;
Query Match 9.5%; Score 124; DB 1; Length 825;
Best Local Similarity 28.3%; Pred. No. 0.4;
Matches 60; Conservative 18; Mismatches 82; Indels 52; Gaps 9;

Qy 6 RKRGAAGAEAAVEVTVGVTRSRSAATGGVAKVAPRKRAPAGE-----PAAA 58
Db 513 RKRGRSDGPAASSASS--SAAPRSLAPQGVAKRAAPR--RAPDSGDRGHGLAP 568
Qy 59 VSAGDGGSCVHILSRMLFMAPQOPQPSVPTPVEAAGCAQQAALAAGLSRCSS 118
Db 569 ASAGA-----APPASPS-----SQAAVAASSSSASSSSASSSSAS 605
Qy 119 TASSVNLGLGQSGHTCRSYDAAPAGDHLVDVSAASNSGSGPDRRRTTPSSRAHG 178
Db 606 SSSASSSSASSSSASSSSASSAGAG-----SVASAGAG---ERRTSLGPRAAA 655
Qy 179 ELSDESLAGHKT-----GFSLPATPAEL 205
Db 656 PRGPRK---CAKTRHABGGPEPGARDPAPGL 684

RESULT 2
PGBM HUMAN
ID PGBM HUMAN STANDARD; PRT; 4391 AA.
AC P98150; Q16287; Q9H3V5;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
GN HSPG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112994; PubMed=1730768;
RA Kallunki P., Tryggvason K.;
RT "Human basement membrane heparan sulfate proteoglycan core protein: a
RT 467-kD protein containing multiple domains resembling elements of the
RT low density lipoprotein receptor, laminin, neural cell adhesion
RT molecules, and epidermal growth factor.";
RL J. Cell Biol. 116:559-571(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Skin;
RX MEDLINE=92235084; PubMed=1569102;
RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT "Primary structure of the human heparan sulfate proteoglycan from
RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT domains homologous to the low density lipoprotein receptor, laminin,
RT neural cell adhesion molecules, and epidermal growth factor.";
RL J. Biol. Chem. 267:8544-8557(1992).
RN [3]
RP SEQUENCE OF 22-4391 FROM N.A., AND VARIANT SJS1 TYR-1532.
RX MEDLINE=20553141; PubMed=1101850;
RA Nicole S., Davoine C.-S., Topaloglu H., Cattolico L., Barral D.,
RA Beighton P., Ben-Hamida C., Hammouda H., Cruaud C., White P.S.,
RA Sanson D., Urtizberea J.A., Lehmann-Horn F., Weissenbach J.,
RA Hentati F., Fontaine B.;
RT "Perlecan, the major proteoglycan of basement membranes, is altered in
RT patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia).";
RL Nat. Genet. 26:480-483(2000).
RN [4]
RP SEQUENCE OF 1016-1470 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=91365376; PubMed=1679749;
RA Dodge G.R., Kovalszky I., Chu M.L., Hassell J.R., McBride O.W.,
RA Yi H.F., Iozzo R.V.;
RT "Heparan sulfate proteoglycan of human colon: partial molecular
```

```
RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
RT short arm of human chromosome 1.";
RL Genomics 10:673-680(1991).
RN [5]
RP SEQUENCE OF 890-1396 FROM N.A.
RC TISSUE=Fibrosarcoma;
RX MEDLINE=92120660; PubMed=1685141;
RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
RA Tryggvason K.;
RT "Cloning of human heparan sulfate proteoglycan core protein,
RT assignment of the gene (HSPG2) to lp36.1--p35 and identification of
RT a BamHI restriction fragment length polymorphism.";
RL Genomics 11:389-396(1991).
RN [6]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=94052171; PubMed=8234307;
RA Cohen I.R., Graessel S., Murdoch A.D., Iozzo R.V.;
RT "Structural characterization of the complete human perlecan gene and
RT its promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
RN [7]
RP CARBOHYDRATE-LINKAGE SITE ASN-2121.
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666(2003).
CC -!- FUNCTION: This protein is an integral component of basement
CC membranes. It is responsible for the fixed negative electrostatic
CC charge and is involved in the charge-selective ultrafiltration
CC properties. It serves as an attachment substrate for cells.
CC -!- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
CC dimers or stellate structures. It interacts with other basement
CC membrane components such as laminin, prolargin and collagen type
CC IV.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Found in the basement membranes.
CC -!- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
CC AND O-LINKED OLIGOSACCHARIDES.
CC -!- DISEASE: Defects in HSPG2 are the cause of Schwartz-Jampel
CC syndrome (SJS1) (MIM:255800); a rare autosomal recessive disorder
CC characterized by permanent myotonia (prolonged failure of muscle
CC relaxation) and skeletal dysplasia, resulting in reduced stature,
CC kyphoscoliosis, bowing of the diaphyses and irregular epiphyses.
CC -!- SIMILARITY: Contains 11 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 11 laminin EGF-like domains.
CC -!- SIMILARITY: Contains 3 laminin IV domains.
CC -!- SIMILARITY: Contains 22 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 3 laminin G-like domains.
CC -!- SIMILARITY: Contains 4 EGF-like domains.
CC -!- SIMILARITY: Contains 1 SEA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62515; CAA44373.1; -
CC EMBL; M85289; AAA52700.1; -
CC EMBL; AL445795; CAC18534.1; -
CC EMBL; M64283; AAA52699.1; -
CC EMBL; S76436; AAB21121.2; -
CC EMBL; I22078; -; NOT_ANNOTATED_CDS.
CC F01; A38096; A38096.
CC HSPG; F00740; IEDM.
CC Sienaa-2DPAGE; P98160; -
CC Genew; HGNC:5273; HSPG2.
CC MIM; 142461; -
CC MIM; 255800; -
CC InterPro; IPR008985; ConA_like_lec_gl.
```



```

DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig c2.
DR InterPro; IPR003596; Ig v.
DR InterPro; IPR000034; Laminin B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000082; SEA_domain.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00047; Ig; 22.
DR Pfam; PF00052; laminin_B; 3.
DR Pfam; PF00053; laminin_EGF; 7.
DR Pfam; PF00054; laminin_G; 3.
DR Pfam; PF00057; ldl_recept_a; 4.
DR Pfam; PF01390; SEA; 1.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRODOM; PD003031; Laminin_B; 3.
DR SMART; SM00181; EGF; 15.
DR SMART; SM00180; EGF_Lam; 12.
DR SMART; SM00409; IG; 22.
DR SMART; SM00408; IGc2; 21.
DR SMART; SM00406; IGV; 7.
DR SMART; SM00281; LamB; 3.
DR SMART; SM00282; LamG; 3.
DR SMART; SM00192; LDLA; 4.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00022; EGF 1; 9.
DR PROSITE; PS01186; EGF 2; 6.
DR PROSITE; PS00026; EGF 3; 4.
DR PROSITE; PS00835; IG_LIKE; 22.
DR PROSITE; PS00025; LAM_G_DOMAIN; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR PROSITE; PS01209; LDLRA 1; 4.
DR PROSITE; PS00068; LDLRA 2; 4.
DR PROSITE; PS00024; SEA; 1.
DR Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
KW Extracellular matrix; EGF-like domain; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 4391
FT FT
FT DOMAIN 80 194
FT DOMAIN 198 235
FT DOMAIN 284 320
FT DOMAIN 324 360
FT DOMAIN 367 404
FT DOMAIN 405 504
FT DOMAIN 521 530
FT DOMAIN 531 730
FT DOMAIN 731 763
FT DOMAIN 814 871
FT DOMAIN 879 923
FT DOMAIN 924 933
FT DOMAIN 934 1125
FT DOMAIN 1126 1158
FT DOMAIN 1159 1208
FT DOMAIN 1209 1265
FT DOMAIN 1275 1324
FT DOMAIN 1325 1345
FT DOMAIN 1345 1529
FT DOMAIN 1530 1562
FT DOMAIN 1563 1612
FT DOMAIN 1613 1670
FT DOMAIN 1677 1771
FT DOMAIN 1772 1865
FT DOMAIN 1866 1955
FT DOMAIN 1956 2051
FT DOMAIN 2052 2151
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig c2.
DR InterPro; IPR003596; Ig v.
DR InterPro; IPR000034; Laminin B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000082; SEA_domain.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00047; Ig; 22.
DR Pfam; PF00052; laminin_B; 3.
DR Pfam; PF00053; laminin_EGF; 7.
DR Pfam; PF00054; laminin_G; 3.
DR Pfam; PF00057; ldl_recept_a; 4.
DR Pfam; PF01390; SEA; 1.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRODOM; PD003031; Laminin_B; 3.
DR SMART; SM00181; EGF; 15.
DR SMART; SM00180; EGF_Lam; 12.
DR SMART; SM00409; IG; 22.
DR SMART; SM00408; IGc2; 21.
DR SMART; SM00406; IGV; 7.
DR SMART; SM00281; LamB; 3.
DR SMART; SM00282; LamG; 3.
DR SMART; SM00192; LDLA; 4.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00022; EGF 1; 9.
DR PROSITE; PS01186; EGF 2; 6.
DR PROSITE; PS00026; EGF 3; 4.
DR PROSITE; PS00835; IG_LIKE; 22.
DR PROSITE; PS00025; LAM_G_DOMAIN; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR PROSITE; PS01209; LDLRA 1; 4.
DR PROSITE; PS00068; LDLRA 2; 4.
DR PROSITE; PS00024; SEA; 1.
DR Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
KW Extracellular matrix; EGF-like domain; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 4391
FT FT
FT DOMAIN 80 194
FT DOMAIN 198 235
FT DOMAIN 284 320
FT DOMAIN 324 360
FT DOMAIN 367 404
FT DOMAIN 405 504
FT DOMAIN 521 530
FT DOMAIN 531 730
FT DOMAIN 731 763
FT DOMAIN 814 871
FT DOMAIN 879 923
FT DOMAIN 924 933
FT DOMAIN 934 1125
FT DOMAIN 1126 1158
FT DOMAIN 1159 1208
FT DOMAIN 1209 1265
FT DOMAIN 1275 1324
FT DOMAIN 1325 1345
FT DOMAIN 1345 1529
FT DOMAIN 1530 1562
FT DOMAIN 1563 1612
FT DOMAIN 1613 1670
FT DOMAIN 1677 1771
FT DOMAIN 1772 1865
FT DOMAIN 1866 1955
FT DOMAIN 1956 2051
FT DOMAIN 2052 2151

```

```

FT DOMAIN 2152 2244 IG-LIKE C2-TYPE 7.
FT DOMAIN 2245 2340 IG-LIKE C2-TYPE 8.
FT DOMAIN 2341 2436 IG-LIKE C2-TYPE 9.
FT DOMAIN 2437 2533 IG-LIKE C2-TYPE 10.
FT DOMAIN 2534 2629 IG-LIKE C2-TYPE 11.
FT DOMAIN 2630 2726 IG-LIKE C2-TYPE 12.

Query Match 8.7%; Score 113; DB 1; Length 4391;
Best Local Similarity 25.5%; Pred. No. 12;
Matches 75; Conservative 27; Mismatches 83; Indels 98; Gaps 16;

QY 2 GKNMKRCGAGAEVAAVEVT-----QVVG--VRTRSGAAATGG-----VA 41
DB 2499 GEYVCRVVGSGTQFASVLTIQRLSGSHSQGVAYPVRISSASLANGHTLDLNCUVA 2558

QY 42 KVAPR-----RKEAP-----ACEPAAAVSAG----- 62
DB 2559 SQAPHTITWYKRGSLPSRHQIVGSLRAIPQVTPADSGEYVCHVNSGAGSRETSLIVTIQ 2618

QY 63 GDGSGCIYHLRSRMLFMAPPQPSVDSPVPEA----ADGAAGQGAALAAAGLSRCSS 118
DB 2619 GSGSS---HVPS---VSPIRIES--SSPTVEGQTLDLNCVVARQQAIIITWKRGSS 2669

QY 119 TIAS-----SVNLIGL-----GQGSHTCR---SYDAAEAGGDHVLVDVSAANSNGSGPDRE 166
DB 2670 LPSRHQTHGSHLRHQMVSADSGEYVCRANNIDALEAS---IVLSVPSFAGSPSPAGSS 2726

QY 167 ---RRETTSPSSRAHGELSDLESDLAG-----HKTGPSLPA 198
DB 2727 MPRIESSSHVAEGETLDLNCVVPQAHQAQVTHWKRGSLPS 2769

RESULT 3
IE18_PRVKA STANDARD; PRT; 1446 AA.
ID IE18_PRVKA STANDARD; PRT; 1446 AA.
AC P33479;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Immediate-early protein IE180.
GN IE.
OS Pseudorabies virus (strain Kaplan) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OC NCBI_TaxID=33703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021039; PubMed=2171211;
RA Vileck C., Kozmik Z., Paces V., Schirm S., Schwyzler M.;
RT "Pseudorabies virus immediate-early gene overlaps with an oppositely
RT oriented open reading frame: characterization of their promoter and
RT enhancer regions.";
RL Virology 179:365-377 (1990).
CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -!- PTM: A long stretch of serine residues may be a major site of
CC phosphorylation.
CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: M34651; AAA47470.1; -.
DR PIR; A45344; A45344.
DR InterPro; IPR005205; Herpes_ICP4_C.
DR InterPro; IPR005206; Herpes_ICP4_N.

```

DR Pfam; PF03585; Herpes ICP4 C; 1.
DR Pfam; PF03584; Herpes ICP4 N; 1.
KW Early protein; Transcription regulation; Trans-acting factor;
KW DNA-binding; Phosphorylation; Nuclear protein.
FT DOMAIN 347 354
FT DOMAIN 379 397
SQ SEQUENCE 1446 AA; 148640 MW; 81P43A3D3DDA068 CRC64;

Query Match 8.6%; Score 112; DB 1; Length 1446;
Best Local Similarity 26.8%; Pred. No. 4.4;
Matches 64; Conservative 19; Mismatches 78; Indels 78; Gaps 11;

QY 9 RGAAGAEEVAVEVT-----QVGVTRSR-----AAATGGVAKVAPRRKAPAGPAA 58
Db 745 RQCAADSVAVARTVAPLVYSDGARAREAAWYAAALFAPANVAARLAEAAAPGPA 804

QY 59 VSAAGDGGSCYIHLRSMLFMAPPQPSVDSVPTVEAADGA-AGQOAGLAALAGLRCS 117
Db 805 EPAPG-----LPLMPEQGLVVPAPAPAAAGAPSLPGSGFSSPASTKS 849

QY 118 S-----TASSVNLGLGQSGHSCRSYDAEAGDHLVDVSAASNGSGFD-----RERR 168
Db 850 SSSTKSSSTKSLGSGSG-----YASSPAAGFPDPAPEPKKK 887

QY 169 ETPSSRAHGLSD-----LESDLAGHK-----TGP-----SI-----PAATPAELI 206
Db 888 REAPCARPFGDEEGLSGALRGDGHGRDDEDRGRRKRRLGLGPDAPDPAPALL 946

RESULT 4
TBC2 CHLRE STANDARD; PRT; 1115 AA.
ID TBC2 CHLRE STANDARD; PRT; 1115 AA.
AC Q8VXF3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tbc2 translation factor, chloroplast precursor.
GN TBC2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=137C / CC-125;
RA MEDLINE=22053252; PubMed=12045185;
RA Auchincloss A.H., Zerges W., Perron K., Girard-Bascou J.,
Rochaix J.-D.;
RT "Characterization of Tbc2, a nucleus-encoded factor specifically
required for translation of the chloroplast psbc mRNA in
Chlamydomonas reinhardtii";
RL J. Cell Biol. 157:953-962(2002).
CC -!- FUNCTION: Required for expression of the chloroplast encoded psbc
mRNA, most likely for translation initiation. Interacts with the
5'-UTR of psbc.
CC -!- SUBUNIT: Part of a 400 kDa complex which is not stably associated
with RNA.
CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -!- SIMILARITY: IN THE CENTRAL SECTION, TO CRP1 OF ZEA MAYS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ427966; CAD20887.1; -
DR Chloroplast; Transit peptide; Repeat.
KW TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
FT CHAIN ? 1115 TBC2 TRANSLATION FACTOR.
FT DOMAIN 483 1068 9 X 38 AA APPROXIMATE REPEATS.

FT REPEAT 483 521 1.
FT REPEAT 607 645 2.
FT REPEAT 685 723 3.
FT REPEAT 724 763 4.
FT REPEAT 764 803 5.
FT REPEAT 804 842 6.
FT REPEAT 843 880 7.
FT REPEAT 990 1029 8.
FT REPEAT 1030 1068 9.
FT DOMAIN 187 210 SER-RICH.
FT DOMAIN 253 256 POLY-ASP.
FT DOMAIN 271 326 ALA-RICH.
FT DOMAIN 370 377 POLY-GLN.
SQ SEQUENCE 1115 AA; 114823 MW; 2239799E91C5D8F7 CRC64;

Query Match 8.6%; Score 111.5; DB 1; Length 1115;
Best Local Similarity 27.5%; Pred. No. 3.6;
Matches 70; Conservative 28; Mismatches 104; Indels 53; Gaps 11;

QY 30 RSRSAATGGVAK---VAPRRKRAPAGPAAAVSAGDGGSCY-IHLRSMLFM-----79
Db 164 RAGSGASTSGRARGWGGSGPGRN---GSGSSSVVNGSSSSSSSSSLAMGMQUSM 219

QY 80 -----APPOQPSVDS-----VPTVEAADGAAGQO-----GAALAAAGLRCS 118
Db 220 ASIGDDVVGVNAGVPVPSGGADALLDLEMSSILDDDGAGARQLQMSDDDLAAGLEAAT 279

QY 119 TASSVNLGLGQSGHSCRSYDAEAGDHLV-----DVSAASNGSGPDRPRRTTFS 173
Db 280 TTAAPAGVAAAGTGAGAAADAASSAPSLVAAAAAASAPSSPDVARTLTL 339

QY 174 SRAGELSDLESDLAGHKTGSLPAA-TPAAELIVPAHEIQEFFAAAACAKEFASKY 232
Db 340 SRAFS--LGLD-----LSGPQLAAVFTGLAVLRPRQOQOQOQOQAGAGAGAGG- 391
QY 233 NFDFVRGVPLDAGGR 247
Db 392 ----VGGVGSAGDR 402

RESULT 5
HLES DROME
ID HLES DROME STANDARD; PRT; 1077 AA.
AC Q02308; Q9VDK0; Q9VDK1;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hairless protein.
GN H OR CG5460.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=92387549; PubMed=1516831;
RA Bang A.G., Posakony J.W.;
RT "The Drosophila gene Hairless encodes a novel basic protein that
controls alternative cell fates in adult sensory organ development.";
RL Genes Dev. 6:1752-1769(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=93041287; PubMed=1419850;
RA Maier D., Stumm G., Kuhn K., Preiss A.;
RT "Hairless, a Drosophila gene involved in neural development, encodes
a novel, serine rich protein.";
RL Mech. Dev. 38:143-156(1992).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;

RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Mocomura K.,
RA Nakano G., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Samei G., Seki Y., Sivasubramaniam S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.,
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.",
RL DNA Res. 3:363-377(1996).
CC -!- SIMILARITY: Belongs to the tail fiber family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000234; AAC74454.1; ALT_INIT.
DR EMBL; D90774; BAA14966.1; -.
DR EMBL; D90775; BAA14975.1; -.
DR PIR; G64887; G64887.
DR Ecogene; EGI3370; stfR.
DR InterPro; IPR008969; Carboxypeptid_reg.
DR InterPro; IPR005003; Phage_fiber.
DR InterPro; IPR005068; Phage_fiber_2.
DR Pfam; PF03335; Phage_fiber_6.
DR Pfam; PF03406; Phage_fiber_2; 1.
DR Hypothetical protein; Fiber protein; Repeat; Complete proteome.
KW SEQUENCE 1120 AA; 113779 MW; 542E59D71EE795B4 CRC64;
SQ
Query Match 8.1%; Score 105.5; DB 1; Length 1120;
Best Local Similarity 22.4%; Pred. No. 8.9;
Matches 56; Conservative 30; Mismatches 113; Indels 51; Gaps 6;
QY 9 RGAAGAEEVAA--VEVTQVGVVTRSRSAATGSAKAPRRK-RAPAGPAAVSAAGDGG 65
Db 237 RDAASKEAKSSETNASSASSASSATAAGNSAKAAKTSETNARSSTTAAGQASAAA 296
QY 66 GSCYIHLRSMFLMAPPQPPQSDVPTVEADGAGGQQAALAGLRCSSTASSVN- 124
Db 297 G-----SKTAAASASAASTASAGQASASATAAGKSAESASAST 336
QY 125 ----LGLGGRGSHGTCRSYDAE-----AGGDHVLVDVSAANSNGSG 162
Db 337 ATKAGEATEQASAAARSASAKTSTNKASETSAESSKTAASASSASASASAS 396
QY 163 PDRERETTPSRAHGELSDLESDLGHKGT-----PSLPAATPAELIVPPAHEIQEFP 217
Db 397 KDEATRQASAAKSSATTASTKATEAAGSATAAQAQSKTAESATRAETAAKRAEDIASV 456
QY 218 AAEEAAQAKR 227
Db 457 ALEDASTTKK 466
RESULT 11
FXCL MOUSE
ID FXCL MOUSE STANDARD; PRT; 553 AA.
AC Q61572; O88409; Q61582;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Forkhead box protein C1 (Forkhead-related protein FKHL7) (Forkhead-
related transcription factor 3) (FEAC-3) (transcription factor FKHL-1)
DE (Mesoderm/mesenchyme forkhead 1) (MF-1).
GN FOXC1 OR FKHL7 OR FEAC3 OR FKHL1 OR MFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=98297351; PubMed=9635428;
RA Xume T., Deng K.Y., Winfrey V., Gould D.B., Walter M.A., Hogan B.L.M.;
RT "The forkhead/winged helix gene Mfl is disrupted in the pleiotropic
mouse mutation congenital hydrocephalus.",
RL Cell 93:985-996(1998).
RN [2]
RP SEQUENCE OF 69-179 FROM N.A.
RC STRAIN=129;
RX MEDLINE=93361500; PubMed=7689224;
RA Kaestner K.H., Lee K.H., Schloendorff J., Hiemisch H.,
RA Monaghan A.P., Schuetz G.;
RT "Six members of the mouse forkhead gene family are developmentally
regulated.",
RL Proc. Natl. Acad. Sci. U.S.A. 90:7628-7631(1993).
RN [3]
RP SEQUENCE OF 71-187 FROM N.A.
RX MEDLINE=93387221; PubMed=8375339;
RA Sasaki H., Hogan B.L.;
RT "Differential expression of multiple fork head related genes during
gastrulation and axial pattern formation in the mouse embryo.",
RL Development 118:47-59(1993).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed in many embryonic tissues, including
prechondrogenic mesenchyme, pericocular mesenchyme, meninges,
endothelial cells, and kidney.
CC -!- DEVELOPMENTAL STAGE: Expressed during embryogenesis.
CC -!- SIMILARITY: Contains 1 fork-head domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF045017; AAC24209.1; -.
DR EMBL; LI0406; AAA03159.1; -.
DR EMBL; X71939; CAA50741.1; -.
DR PIR; I49674; I49674.
DR HSP; Q63245; 2HFH.
DR TRANSFAC; T02426; -.
DR MGI; MGI:1347466; Foxc1.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00339; FORK_HEAD_3; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DOMAIN 28 33
FT DNA BIND 77 168
FT DOMAIN 169 173
FT DOMAIN 194 197
FT DOMAIN 264 274
FT DOMAIN 375 386
FT DOMAIN 444 451
FT DOMAIN 453 456
FT DOMAIN 486 496
FT CONFLICT 180 186 VKDEEK -> KKEITFI (IN REF. 3).
SQ SEQUENCE 553 AA; 56953 MW; 3CDD12F69CA4F217 CRC64;
Query Match 8.1%; Score 105; DB 1; Length 553;
Best Local Similarity 22.9%; Pred. No. 4.6;
Matches 63; Conservative 33; Mismatches 103; Indels 76; Gaps 13;
QY 31 SRSAATGGVAKVAPRRKAPAGEPAAVSAAGDGGSCYVHLSRMLFMA----PPQPOP 86
Db 243 SPAAALGSGSAATVPKIESPDSS--SSLSGSGSPGSLPSARPLSLDAEAPPPQAP 300
QY 87 -----SVDSVPTPVEAA-DGAAGQQAALAGLRCSSTASSVNLGLG----GQRG- 132

Qy	6	KRCRGAGAE-----VAAVEVTQVGVRTRSAAATGV-----AKVAPRRKRA	50
Dd	23	KKAEAGTBEETGQKSEPOAAAADATEVKESABEKPDAADGEAKAEKEADKAAAKEA	82
Qy	51	PAGEPAAAVSAGDGGSCYIHLRSMFLFAPPQPQPSVDSPVTPVE--AADG-AAQGQA	107
Dd	83	PKAEPEKSEGA-----AEOQEPA---PAPEOAAAPGPAAG--GE	118
Qy	108	ALAAGLSRCSTASSVNLGIGCGRGSHTRCYSDAAEAGGDHVLVDVSSANGSGPDPR	167
Dd	119	APKAG---EASAEASTCAADGAPQEBGEAKKTAPAAGPEAKSDAAPAADS-----KPS	169
Qy	168	RETTPSRAHELGSLDLGHKTGPSLPAAATPAALIVPPAHE	212
Dd	170	TEPAPSCKTTPAASEAPSS-AKAAPAPAPAPAPAPAPAPASSEO	213

```

RESULT 15
ID KLTk MOUSE STANDARD; PRT; 888 AA.
AC P08923;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Leukocyte tyrosine kinase receptor precursor (EC 2.7.1.112).
LN LTK.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=BALB/c;
RX MEDLINE=90291994; PubMed=2357970;
RA Bernards A., de la Monte S.;
RT "The ltk receptor tyrosine kinase is expressed in pre-B lymphocytes
RT and cerebral neurons and uses a non-AUG translational initiator.";
RL EMBO J. 9:2279-2287(1990).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=BALB/c;
RX MEDLINE=92115335; PubMed=1662793;
RA Haase V.H., Snijders A.J., Cooke S.M., Teng M.N., Kaul D.,
RA le Beau M.M., Bruns G.A., Bernards A.;
RT "Alternatively spliced ltk mRNA in neurons predicts a receptor with a
RT larger putative extracellular domain.";
RL Oncogene 6:2319-2325(1991).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM A).
RP MEDLINE=88232962; PubMed=2836739;
RX Ben-Neriah Y., Bauskin A.R.;
RT "Leukocytes express a novel gene encoding a putative transmembrane
RT protein-kinase devoid of an extracellular domain.";
```

RL Nature 333:672-676(1988).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D).
 RX MEDLINE=93141274; PubMed=8380920;
 RA Snijders A.J., Haase V.H., Bernards A.;
 RT "Four tissue-specific mouse ltk mRNAs predict tyrosine kinases that
 RT differ upstream of their transmembrane segment.";
 RL Oncogene 8:27-35(1993).
 RN [5]
 RP SUBCELLULAR LOCATION, AND INTERACTION WITH CALNEXIN (ISOFORM A).
 RX MEDLINE=97150899; PubMed=8995435;
 RA Snijders A.J., Ho S.C., Haase V.H., Pillai S., Bernards A.;
 RT "A lymphocyte-specific ltk tyrosine kinase isoform is retained in the
 RT endoplasmic reticulum in association with calnexin.";
 RL J. Biol. Chem. 272:1297-1301(1997).
 CC -!- FUNCTION: THE EXACT FUNCTION OF THIS PROTEIN IS NOT KNOWN, IT IS
 CC PROBABLY A RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBUNIT: Isoform A binds calnexin.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Isoform A and
 CC isoform C are retained in the endoplasmic reticulum.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative promoter;
 CC Comment=2 isoforms, A and B, may be produced by use of
 CC alternative promoters;
 CC Event=Alternative splicing; Named isoforms=4;
 CC Comment=Additional isoforms seem to exist;
 CC Name=B;
 CC IsoId=P08923-1; Sequence=Displayed;
 CC Name=A;
 CC IsoId=P08923-2; Sequence=VSP 002950, VSP 002951, VSP 002952;
 CC Note=May be produced by use of alternative promoters. Starts at
 CC a CUG codon;
 CC Name=B;
 CC IsoId=P08923-3; Sequence=VSP 002950, VSP 002951;
 CC Note=May be produced by use of alternative promoters. Starts at
 CC a CUG codon;
 CC Name=C;
 CC IsoId=P08923-4; Sequence=VSP 002952;
 CC -!- TISSUE SPECIFICITY: Subsets of lymphoid and neuronal cells.
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
 CC receptor subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52621; CAA36848.1; -;
 DR EMBL; X07984; CAA30793.1; ALT INIT.
 DR EMBL; M90470; AAA39451.1; -;
 DR PIR; I58378; I58378.
 DR PIR; S12792; S12792.
 DR HSP; P11362; I5GK.
 DR MGD; MGI:96840; Ltk.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002011; RecepttyrkinsII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00239; RECEPTOR TYR KIN II; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR Transferase; Tyrosine-protein kinase; Transmembrane; Receptor;
 KW Endoplasmic reticulum; ATP-binding; Phosphorylation; Glycoprotein;

KW Signal; Alternative splicing; Alternative promoter usage.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 888 LEUKOCYTE TYROSINE KINASE RECEPTOR.
 FT DOMAIN 17 421 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 422 446 POTENTIAL.
 FT DOMAIN 447 588 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 506 782 PROTEIN KINASE.
 FT NP_BIND 512 520 ATP (BY SIMILARITY).
 FT BINDING 540 540 ATP (BY SIMILARITY).
 FT ACT_SITE 639 639 BY SIMILARITY.
 FT MOD_RES 672 672 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 377 377 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 1 251 Missing (in isoform A and isoform B).
 FT VARSPLIC 252 252 /FTid=VSP 002950.
 FT VARSPLIC 252 252 L->M (in isoform A and isoform B).
 FT VARSPLIC 271 331 /FTid=VSP 002951.
 FT VARSPLIC 271 331 Missing (in isoform A and isoform C).
 FT CONFLICT 789 789 M -> V (IN REF. 3).
 FT CONFLICT 875 875 Q -> H (IN REF. 3).
 SQ SEQUENCE 888 AA; 94436 MW; 3FFCA80BA4863C55 CRC64;
 Query Match 8.0%; Score 104; DB 1; Length 888;
 Best Local Similarity 28.0%; Pred. No. 8.8;
 Matches 60; Conservative 16; Mismatches 94; Indels 44; Gaps 11;
 Qy 8 CRGAAGAEVAAVEVTQVGVTRSRSAATCGVAKVAPRRKAPAGEPAAVSAGDGGG 67
 Db 179 CLGESGEHATTYGTERTIPGWRWAGGGGGGATSF----RLRAGEPEPLVVAAGGGR 234
 Qy 68 CYIHLRMLFMAPPQPPQPSVDSPVTPVE---AADGAAGQOQGAALAGLSCSSTASSVN 124
 Db 235 SY---RRR-----PDRGRTQAVPERLETAAPAGSGRGGA--AGGSGGWTSAHPQ 282
 Qy 125 LGLG---GQSGHTCRSYDA---AERGGDHVLVDVSAANSNGSGGPDRETTPTSSRAH 177
 Db 283 AGRSPREGAEGEGCAEAWAALRWAAAGGFGGGGCAACAAGGGGGG-----YRG 330
 Qy 178 GELSDLESLD---AGHKTGPSLPAAATPAALIVPP 209
 Db 331 GDTS--ESDLLWADGEDGTGF--VHPSGELYLPQ 360
 Search completed: October 1, 2004, 16:51:44
 Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:45:36 ; Search time 116 Seconds
(without alignments)
696.316 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 1304

Sequence: 1 MGKYMRCRGAAGAAVEA.....VRGVLDAAGGFEPWAPVSI 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL_25.*

2: sp_archaea.*

3: sp_bacteria.*

4: sp_fungi.*

5: sp_human.*

6: sp_invertebrate.*

7: sp_mammal.*

8: sp_mhc.*

9: sp_organelle.*

10: sp_phase.*

11: sp_plant.*

12: sp_rodent.*

13: sp_virus.*

14: sp_vertebrate.*

15: sp_unclassified.*

16: sp_rivirus.*

17: sp_bacteriap.*

18: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	17.3	242	10 Q9FW65	Q9FW65 oryza sativ
2	225	17.3	242	10 Q7XDH8	Q7XDH8 oryza sativ
3	212.5	16.3	222	10 Q9FKB5	Q9FKB5 arabidopsis
4	207	15.9	185	10 Q8GT28	Q8GT28 lycopersico
5	179.5	13.8	286	10 Q94CM0	Q94CM0 arabidopsis
6	178	13.7	196	10 Q48597	Q48597 chenopodium
7	178	13.7	289	10 Q48846	Q48846 arabidopsis
8	177.5	13.6	286	10 Q8GYJ3	Q8GYJ3 arabidopsis
9	170	13.0	192	10 Q9FS28	Q9FS28 pisum sativ
10	169.5	13.0	210	10 Q8GT29	Q8GT29 lycopersico
11	163	12.5	207	10 Q8GUA2	Q8GUA2 nicotiana t
12	153	11.7	189	10 Q9LRV0	Q9LRV0 arabidopsis
13	152	11.7	163	10 Q93V92	Q93V92 nicotiana t
14	134	10.3	196	10 Q9LJL5	Q9LJL5 arabidopsis
15	129.5	9.9	204	10 Q8S9U7	Q8S9U7 oryza sativ
16	129.5	9.9	1156	16 Q9Z5A4	Q9Z5A4 streptomyce

17	128	9.8	801	5	Q23635
18	126	9.7	1080	10	Q9SDC1
19	124	9.5	826	12	P89473
20	121.5	9.3	195	10	Q9FX90
21	121.5	9.3	195	10	Q94CL9
22	121.5	9.3	1151	13	O57580
23	121.5	9.3	1340	16	Q9L1H8
24	121.5	9.3	2639	5	O76786
25	121	9.3	191	10	O82809
26	121	9.3	191	10	O8LDX1
27	121	9.3	191	10	O04154
28	121	9.3	2936	5	Q9NKP7
29	119	9.1	1787	10	Q9M4X9
30	118.5	9.1	775	16	Q9F342
31	117.5	9.0	1468	5	Q9GUB5
32	116.5	8.9	452	16	Q9KY45
33	116.5	8.9	605	2	Q9AH41
34	116.5	8.9	1387	16	Q82C89
35	116.5	8.9	2655	5	Q964F4
36	116	8.9	362	10	Q8S031
37	115.5	8.9	1338	4	Q9F615
38	115	8.8	156	10	Q93YF6
39	115	8.8	1414	5	Q26634
40	114.5	8.8	812	16	Q8ZQ81
41	114.5	8.8	832	16	Q9KZY7
42	113.5	8.7	467	16	O06256
43	113.5	8.7	467	16	Q7TW11
44	113	8.7	1541	5	O15837
45	112.5	8.6	3012	5	O97205

ALIGNMENTS

RESULT 1

Q9FW65	PRELIMINARY;	PRT;	242 AA.
ID	Q9FW65		
AC	Q9FW65		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Putative cyclin-dependent kinase inhibitor.		
GN	OSJNB0094K03.16.		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae; Oryzeae; Oryza.		
OX	NCBI TaxID=4530;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Nipponbare;		
RA	Buell C.R., Yuan O., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,		
RA	Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,		
RA	Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,		
RA	Quackenbush J., White O., Salzberg S.L., Fraser C.M.;		
RT	"Oryza sativa chromosome 10 BAC OSJNB0094K03 genomic sequence.";		
RL	Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; AC069145; AAG16867.1; -		
DR	Gramene; Q9FW65; -		
DR	GO; GO:0005634; C:nucleus; IEA.		
DR	GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . . ; IEA.		
DR	GO; GO:0016301; F:kinase activity; IEA.		
DR	GO; GO:0007050; P:cell cycle arrest; IEA.		
DR	InterPro; IPR003175; CDI.		
DR	Pfam; PF02234; CDI; 1.		
KW	Cyclin; Kinase.		
SQ	SEQUENCE 242 AA; 26750 MW; F215A01234735928 CRC64;		

Query Match 17.3%; Score 225; DB 10; Length 242;

Best Local Similarity 28.0%; Pred. No. 5.9e-08;

Matches 84; Conservative 24; Mismatches 80; Indels 112; Gaps 12;

QY 1 MGKYMRCRGAAGAAVEA...VGVTRSRSAATGGVAKVAPRRKRAPAGEPA 56

```
Db 1 MGKYMRAKVVSVGEVAAVMAAALPLGVRTRASLAL-----QKR----- 43
QY 57 AAVSAGDGGGCVIHLRSMLFMAPPQOPQSDVSVPTFVEAADGAGQQAALAGLSRC 116
Db 44 -----QGGE-----YLELRSSRLKLPPLPP-----PPRRRATAAATAATAA----- 83
QY 117 SSTASSVNLGLGQSGHSCRSYDAEA-----GGDHVL-----VDVSAASNGS 161
Db 84 -----ESAEVSVFGENVLELEAMERIAFPVSLNCVKNVTV 120
QY 162 GP-----DRER-----RETPSS-----RAHGELSLSLESLAGH 190
Db 121 APSKYPVGCATCLOGTKGVPERDEQVLRHFWNTRETTPCSLIRDPDTISTPGSTTRS 180
QY 191 KTGPSLPAATPAELIIVPPAHEIQEFPAAAEAAQAKRFASKYNFDFVRGVPLDAGGFWE 250
Db 181 HSSSHCKVQTPVRHNIIPASAELEAFPAEQRQAFIDKYNFDPVNDCLPLP--GRFEW 238
```

RESULT 2

```
Q7XDH8 PRELIMINARY; PRT; 242 AA.
```

```
AC Q7XDH8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative cyclin-dependent kinase inhibitor.
GN OSUNB0094K03.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA The Rice Chromosome 10 Sequencing Consortium;
RT "in-depth view of structure, activity, and evolution of rice
  chromosome 10.";
RL Science 300:1566-1569(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017104; AAP54233.1; -.
KW Cyclin; Kinase.
SQ SEQUENCE 242 AA; 26750 MW; F215A01234735928 CRC64;
```

```
Query Match 17.3%; Score 225; DB 10; Length 242;
Best Local Similarity 28.0%; Pred. No. 5.9e-08;
Matches 84; Conservative 24; Mismatches 80; Indels 112; Gaps 12;

QY 1 MGKYMRAKVVSVGEVAAVMAAALPLGVRTRASLAL-----QKR----- 43
Db 1 MGKYMRAKVVSVGEVAAVMAAALPLGVRTRASLAL-----QKR----- 43
QY 57 AAVSAGDGGGCVIHLRSMLFMAPPQOPQSDVSVPTFVEAADGAGQQAALAGLSRC 116
Db 44 -----QGGE-----YLELRSSRLKLPPLPP-----PPRRRATAAATAATAA----- 83
QY 117 SSTASSVNLGLGQSGHSCRSYDAEA-----GGDHVL-----VDVSAASNGS 161
Db 84 -----ESAEVSVFGENVLELEAMERIAFPVSLNCVKNVTV 120
QY 162 GP-----DRER-----RETPSS-----RAHGELSLSLESLAGH 190
Db 121 APSKYPVGCATCLOGTKGVPERDEQVLRHFWNTRETTPCSLIRDPDTISTPGSTTRS 180
QY 191 KTGPSLPAATPAELIIVPPAHEIQEFPAAAEAAQAKRFASKYNFDFVRGVPLDAGGFWE 250
Db 181 HSSSHCKVQTPVRHNIIPASAELEAFPAEQRQAFIDKYNFDPVNDCLPLP--GRFEW 238
```

RESULT 3

```
Q9FKB5 PRELIMINARY; PRT; 222 AA.
AC Q9FKB5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Genomic DNA, chromosome 5, TAC clone:K24G6 (Cyclin-dependent kinase
  inhibitor 3).
GN KRP3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
  Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
  Sequence features of the regions of 1,367,185 bp covered by 19
  DNA Res. 5:203-216(1998).";
RL DNA Res. 5:203-216(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=21342510; PubMed=11449057;
RA de Veylder L., Beeckman T., Beeckman G.T.S., Krols L., Terras F.,
  Landrieu I., Van Der Schueren E., Maes S., Naudts M., Inze D.;
RT "Functional analysis of Cyclin-dependent kinase inhibitors of
  Arabidopsis.";
RL Plant Cell 13:1653-1668(2001).
DR EMBL; AB012242; BAB09435.1; -.
DR EMBL; AJ301554; CAC41617.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a.; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase.
SQ SEQUENCE 222 AA; 24925 MW; 7A3EB2C9A29688A7 CRC64;
```

```
Query Match 16.3%; Score 212.5; DB 10; Length 222;
Best Local Similarity 29.6%; Pred. No. 3.9e-07;
Matches 79; Conservative 33; Mismatches 94; Indels 61; Gaps 12;

QY 1 MGKYMRAKVVSVGEVAAVMAAALPLGVRTRASLAL-----QKR----- 43
Db 1 MGKYMRAKVVSVGEVAAVMAAALPLGVRTRASLAL-----QKR----- 43
QY 57 AAVSAGDGGGCVIHLRSMLFMAPPQOPQSDVSVPTFVEAADGAGQQAALAGLSRC 116
Db 44 -----QGGE-----YLELRSSRLKLPPLPP-----PPRRRATAAATAATAA----- 83
QY 117 SSTASSVNLGLGQSGHSCRSYDAEAAGDHVLVDVSAASNGS--GSGPDRERRET 170
Db 91 SRSRVDVSVNS-----VPVAQSNDECECFDNFVSVQVSCGSLGFSRSHSTES 139
QY 171 TPSSRAHGELESDLAGHKTGFSPLPAATP-----AAELIIVPPAHEIQEFPAAAEAAQAK 226
Db 140 TPCN-----FVEDMEIMVTPGSSSTRMCRTKEYTREQDNVPTTSEMEEFFAYARQQOOR 195
QY 227 RPAKYNFDFVRGVPLDAGGFWEAPY 253
Db 196 LFMEKYNFDFVNDIPL--SGRYEWVQV 220
```

RESULT 4

```
QSGT28
```



```

Query Match      13.7%; Score 178; DB 10; Length 196;
Best Local Similarity 27.5%; Pred. No. 7.8e-05;
Matches 61; Conservative 29; Mismatches 102; Indels 30; Gaps 7;

QY 33 SAAATGGVAKVAPRRKAPAGEPAAAVSAGDGGSCYVHLRSMLFMAPQPSVDSVP 92
Db 1 MGKVIKSKIDGAGAGAGGGGGGGGESSIALMDVVPSSSSSLGVLTRAKSLALQOQQ 60
QY 41 AKVAPRRKAPAGEPAAAVSAG-----GDGSCYVHLRSMLFMAP----- 81
Db 61 QRCILQKPSFSLPPTSASPNPSSKQMKKKQOQMNDCGS-YLQLSRRLQKKPPIVVI 119
QY 82 -----POQPSVDSVPTPVEAADGAAGQOQGAALAGLSRCSSTASSV 123
Db 120 RSTYRRKQQRNETCGRNPFRNLSI-----RGDG-----SRSDSVSESV 161
QY 124 NLGLGGQSGSHTCRS-----YDAEAGGDHVLVDVSAASNSGSGPDRRRTTPSSRA 176
Db 162 VFGKDKDLISEINKDPFGQNFPLEE---EHT-----QSFNRITRESTPGSLI 207
QY 177 H-----GELSLESDLA-GHKTGPSLPAATPAABELIYVPPAHEIOEFF 217
Db 208 RRPEIMTTPGSSTKLNCVSESNOQSDLSRSHRRPTTP-----EMDEFF 253
QY 218 AAAEAAQAKRFASKYNFDFVRGVFLDAGRFEWAPV 253
Db 254 SGABEQKQFIEKYNFDPVNEQLP--GRFEWTKV 287

RESULT 7
O48846 PRELIMINARY; PRT; 289 AA.
AC O48846; Q9FR83;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Expressed protein (Cyclin-dependent kinase inhibitor).
GN ATG323710 OR ACK2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Kaul S.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Park S.C., Cho J.W.;
RA "A novel cyclin-dependent kinase inhibitor (ACK2) interacting with G1
RT specific cyclin";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC003974; AAC0492.2; -
DR EMBL; AF123315; AAG41215.1; -
DR PIR; T00797; T00797.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase.
SQ SEQUENCE 289 AA; 32068 MW; BD7F524A308E11CF CRC64;

Query Match      13.7%; Score 178; DB 10; Length 289;
Best Local Similarity 24.1%; Pred. No. 0.00012;
Matches 81; Conservative 34; Mismatches 89; Indels 132; Gaps 15;

QY 1 MGKVMRKCR-----GAAG-----AFVAAVEV-----TQVVGVTGRSRAATGGV 40
Db 1 MGKVIKSKIDGAGAGAGGGGGGGGESSIALMDVVPSSSSSLGVLTRAKSLALQOQQ 60
QY 41 AKVAPRRKAPAGEPAAAVSAG-----GDGSCYVHLRSMLFMAP----- 81
Db 61 QRCILQKPSFSLPPTSASPNPSSKQMKKKQOQMNDCGS-YLQLSRRLQKKPPIVVI 119
QY 82 -----POQPSVDSVPTPVEAADGAAGQOQGAALAGLSRCSSTASSV 123
Db 120 RSTYRRKQQRNETCGRNPFRNLSI-----RGDG-----SRSDSVSESV 161
QY 124 NLGLGGQSGSHTCRS-----YDAEAGGDHVLVDVSAASNSGSGPDRRRTTPSSRA 176
Db 162 VFGKDKDLISEINKDPFGQNFPLEE---EHT-----QSFNRITRESTPGSLI 207
QY 177 H-----GELSLESDLA-GHKTGPSLPAATPAABELIYVPPAHEIOEFF 217
Db 208 RRPEIMTTPGSSTKLNCVSESNOQSDLSRSHRRPTTP-----EMDEFF 253
QY 218 AAAEAAQAKRFASKYNFDFVRGVFLDAGRFEWAPV 253
Db 254 SGABEQKQFIEKYNFDPVNEQLP--GRFEWTKV 287

RESULT 8
O8GVJ3 PRELIMINARY; PRT; 286 AA.
AC O8GVJ3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative f2417 cyclin-dependent kinase inhibitor ACK2.
GN ATG323710.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RA "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK117586; BAC4243.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase; Cyclin.
SQ SEQUENCE 286 AA; 31689 MW; A541B53AB62A12BD CRC64;

Query Match      13.6%; Score 177.5; DB 10; Length 286;
Best Local Similarity 24.1%; Pred. No. 0.00013;
Matches 81; Conservative 33; Mismatches 87; Indels 135; Gaps 15;

QY 1 MGKVMRKCR-----GAAG-----AFVAAVEV-----TQVVGVTGRSRAATGGV 40
Db 1 MGKVIKSKIDGAGAGAGGGGGGGGESSIALMDVVPSSSSSLGVLTRAKSLALQOQQ 60
QY 41 AKVAPRRKAPAGEPAAAVSAG-----GDGSCYVHLRSMLFMAP----- 81
Db 61 QRCILQKPSFSLPPTSASPNPSSKQMKKKQOQMNDCGS-YLQLSRRLQKKPPIVVI 119
QY 82 -----POQPSVDSVPTPVEAADGAAGQOQGAALAGLSRCSSTASSV 123
Db 120 RSTYRRKQQRNETCGRNPFRNLSI-----RGDG-----SRSDSVSESV 161
QY 124 NLGLGGQSGSHTCRS-----YDAEAGGDHVLVDVSAASNSGSGPDRRRTTPSSRA 176
Db 162 VFGKDKDLISEINKDPFGQNFPLEE---EHT-----QSFNRITRESTPGSLI 207
QY 177 H-----GELSLESDLA-GHKTGPSLPAATPAABELIYVPPAHEIOEFF 217
Db 208 RRPEIMTTPGSSTKLNCVSESNOQSDLSRSHRRPTTP-----EMDEFF 253
QY 218 AAAEAAQAKRFASKYNFDFVRGVFLDAGRFEWAPV 253
Db 254 SGABEQKQFIEKYNFDPVNEQLP--GRFEWTKV 287

```

```
QY 124 NLGLGGQSGSHTCRS-----YDAAGAGDHVLVDVSAASNSGSGPDRERRETTSSRA 176
Db 162 VFGKDKDLISEINKPTFGQNFPLEE---EHT-----QRTRESTPCSLI 204
QY 177 H-----GELSDLESDLA-GHKTGPSLPAAPAAELIVPPAHETQEFF 217
Db 205 RREIMTTPGSKTLNVCSESNQREDLSRSHRRPTTP-----ENGEFF 250
QY 218 AAAAQAQKRFASKYNFDFVRGVPLDAGGRFEWAPV 253
Db 251 SGAEEOQKQFIEKYNFDPVNEQLP--GRFEWTKV 284

RESULT 9
Q9FS28 ID Q9FS28 PRELIMINARY; PRT; 192 AA.
AC Q9FS28;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cyclin dependent kinase inhibitor.
GN CKI.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3889;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Axillary bud;
RA Shimizu-Sato S., Mori H.;
RT "A cDNA from Pisum sativum encoding the cyclin-dependent kinase
inhibitor (CKI) homologue."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029483; BAB20860.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW KINASE.
SQ SEQUENCE 192 AA; 21921 MW; E76D734000885EB2 CRC64;

Query Match 13.0%; Score 170; DB 10; Length 192;
Best Local Similarity 27.8%; Pred. No. 0.00027;
Matches 65; Conservative 28; Mismatches 81; Indels 60; Gaps 11;

QY 25 VGVTRRSRA---AATGGVAKVAPRRKRAPAGAPAAASAGSGGSCYIHLRSLFMFAP 81
Db 6 VGVTRARAALAMEATSSPRTTKRK-----INRTENRKFT- 43
QY 82 PGPQSVSV-PTPVEADGAAGQAALAGLSCSSSTASSVNLGLGQSGSHTCRSYD 140
Db 44 --VKPKIATVREPTVTEKHSGSTSDSEFPAS---CCSSNGSVEL-----DEERIKSLD 92
QY 141 AAAGAGDHVLVDVSA--SSNSGSPDRERRETTSSRAHGLSDLSLGHKTGPSL 196
Db 93 -----LEVASQGTSTCNDEELEREMSKSEFRGNSHELES---METNSRR 138
QY 197 PAATPAELIVPPAHETQEFFAAAAQAQKRFASKYNFDFVRGVPLDAGGRFEW 250
Db 139 PISSPKK---TTEVELEBEFFAAAEKIQKQKYNVYDILKDVPLE--GRFEW 187

RESULT 10
Q8GT29 ID Q8GT29 PRELIMINARY; PRT; 210 AA.
AC Q8GT29;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE P27KIP1-related-protein 1.
GN K3P1
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wv106; TISSUE=Locular;
RA Bisbis B., Joubes J., Hernould M., Inze D., Raymond P., Chevalier C.;
RT "Cyclin-dependent kinase inhibitors are highly expressed in
endoreduplicating tissues during the early development of tomato
fruits."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ441249; CAD29648.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
SQ SEQUENCE 210 AA; 23845 MW; 38FB15C35044FED6 CRC64;

Query Match 13.0%; Score 169.5; DB 10; Length 210;
Best Local Similarity 26.4%; Pred. No. 0.00032;
Matches 70; Conservative 23; Mismatches 103; Indels 69; Gaps 8;

QY 1 MCKYMKRCRGAAGAEVAVVTVGVVTRTSRAATGGVAKVAPRRKRAPAGAPAAVS 60
Db 1 MGKYIRKTR-----KTEDVSPGLVLRALNALN----- 29
QY 61 AGDGGSCYIHLRSLFMFAPQPOPSVDSVPTPVEADGAAGQAALAGLSCSSTA 120
Db 30 -GGDGS-YLELRRLV-----KPTVLEGRQKNGVKNPILVNFNPQ 73
QY 121 SSVNGLGGQSGSHTCRSYDAAGDHVLVDVSAASN--SGSGPDRERRETTSSRAHG 178
Db 74 QIPNVCSNEEGKGVKEMENQEKESCLGPDSCFENLEPEGRKRTTRESTPCS---- 129
QY 179 ELSDSLGLAGHKTGPSLPAATPAEL-----IVPPAHETQEFFAAAAQAQKRF 228
Db 130 ----LIRSDNIQTGSSSTRRTNANEANGRVNSIQPTIPTLEMEEFFTRAKEQQRKF 185
QY 229 ASKYNFDFVRGVPLDAGGRFEWAPV 253
Db 186 IEKYNFDPVNEKPLP--GRFEWTKV 208

RESULT 11
Q8GUA2 ID Q8GUA2 PRELIMINARY; PRT; 207 AA.
AC Q8GUA2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cyclin-dependent kinase inhibitor.
GN KIS2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Jasinski S., Saraiva Leite C., Domenichini S., Stevens R., Raynaud C.,
RA Perennes C., Bergounioux C., Glab N.;
RT "NtKIS2, a novel tobacco cyclin-dependent kinase inhibitor
differentially expressed during the cell cycle and plant
development."
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ517189; CAD56868.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
```

```

DR GO:0016301; F:kinase activity; IEA.
DR GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase.
SQ SEQUENCE 207 AA; 23278 MW; F5AE70AD2261BC2D CRC64;

Query Match 12.5%; Score 163; DB 10; Length 207;
Best Local Similarity 26.4%; Pred. No. 0.00088;
Matches 69; Conservative 40; Mismatches 80; Indels 72; Gaps 13;

QY 1 MGKTMKRCGAAGAEEVAEVTVQVGVTRTSRSAAATGGVAKVAPRRKRAPAGE-----54
DB 1 MRRYKSKSGI--GEMTMEVGKVEEVAVK-MTASAEVLEVAHADHAKKRCMCDGDLML 57

QY 55 PAAVAGSGGGGCGYIHLRMLFMAPPQPSVDSVPTVEAADGAAGQGAALAGL-113
DB 58 PTVA-----C-VRSHGSDILVAQ-----ESLVTPT-TSELSPKENAALSSDFD 98

QY 114 -----SRCSSTASSVNLGLGGGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRRE 169
DB 99 DALASCCASNGSSKSLD-----EETVEI-ATSKSKESIKABLRQ 138

QY 170 TTPSSRAHGLSDLESGLAGHTGSLPAATPAELIYPPAHEIOEFPFAAAQAARFA 229
DB 139 MEPTTRAHHPS-----RRRLTEEKMPSET-----DLEEFFAAAEKDKLKFT 181

QY 230 SKYNDFVRGVLDPAGRFEW 250
DB 182 KKNFDFVKEEPLF--GHYEW 200

RESULT 12
Q9LR10 PRELIMINARY; PRT; 189 AA.
AC Q9LR10
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similarity to cyclin-dependent kinase inhibitor protein (Cyclin-
DE dependent kinase inhibitor 5).
GN KRP5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucotsid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Nakamura Y.;
RL MEDLINE=20277480; PubMed=10819329;
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
[3]
RP SEQUENCE FROM N.A.
RC MEDLINE=21342510; PubMed=11449057;
RA de Veylder L., Beeckman T., Beemster G.T.S., Krols L., Terras F.,
RA Landrieu I., Van Der Schueren E., Maes S., Naudts M., Inze D.;
RT "Functional analysis of Cyclin-dependent kinase inhibitors of
RT Arabidopsis.";
RL Plant Cell 13:1653-1668(2001).
DR EMBL; AB028609; BAB02891.1; -
DR EMBL; AJ301556; CAC41619.1; -
DR GO:0005634; C:nucleus; IEA.
DR GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.

DR GO:0016301; F:kinase activity; IEA.
DR GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Cyclin; Kinase.
SQ SEQUENCE 189 AA; 21424 MW; A66041310E8DDE6 CRC64;

Query Match 11.7%; Score 153; DB 10; Length 189;
Best Local Similarity 25.2%; Pred. No. 0.0038;
Matches 64; Conservative 29; Mismatches 95; Indels 66; Gaps 8;

QY 1 MGKTMKRCGAAGAEEVAEVTVQVGVTRTSRSAAATGGVAKVAPRRKRAPAGEPAAVSA 60
DB 1 MGKTIKSKVAGAVSVKDKGHPFALGPRTRAAA-----KNLALHLRSHSDE-----48

QY 61 AGDGGGCGYIHLRMLFMAPPQPSVDSVPTVEAADGAAGQGAALAGLSRCSSTA 120
DB 49 --ADSFN-YIQLSRRRLVKLP-----LLNTRKQKQQLIPSVNQCOTKN 90

QY 121 SSVNLGLGGGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRRETTPSSRAHGL 180
DB 91 PRASSGPAKKLEPDT-----TTEACGDNERISDCNFGDKGFDLE-----SENRSMI 139

QY 181 SDESGLAGHTGSLPAATPAELIYPPAHEIOEFPFAAAQAARFAASKYNDFVRGV 240
DB 140 SDSKSIQS-----ETEDFFASAEQQQORFFIQKYNDFIVSDN 176

QY 241 PLDAGRGFEWAPVV 254
DB 177 PLP--GRYEWVKVM 188

RESULT 13
Q93V92 PRELIMINARY; PRT; 163 AA.
AC Q93V92
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CDK/cyclin inhibitor.
GN KIS1A OR KIS1.
OS Nicotiana tomentosiformis (Tobacco), and
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4098, 4097;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=N.tomentosiformis, and N.tabacum;
RA Grondard S., Perennes C., Bernardi F., Bergounioux C., Glab N.;
RT "The tobacco CDK inhibitor NtKIS1a is involved in flower
RT development.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297906; CAC82733.1; -
DR EMBL; AJ297904; CAC82731.1; -
DR GO:0005634; C:nucleus; IEA.
DR GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Cyclin.
SQ SEQUENCE 163 AA; 18301 MW; E154A59D491B66B7 CRC64;

Query Match 11.7%; Score 152; DB 10; Length 163;
Best Local Similarity 24.7%; Pred. No. 0.0038;
Matches 63; Conservative 20; Mismatches 76; Indels 96; Gaps 10;

QY 5 MRKCRGAAGAEEVAEVTVQV---VGVTRTSRSAAATGGVAKVAPRRKRAPAGEPAAVSA 61
DB 1 MRKCKGI--EETVMEVSDVDLEVTPTTKKRSISDGDVKLMSPLLR-----46

QY 62 GGDGSGCYIHLRMLFMAPPQPSVDSVPTVEAADGAAGQGAALAGLSRCSSTAS 121

```



```

Db      47  -----CRSH-----|-----|-----SGVGDTPAGSLV-----|:!:SPSS 66
Qy     122  SVNLGLGGQGSHTCRSYDAEAGDHDVLVDVSAAASGSGS-GPDRERRETTPTSSRAHGEL 180
         |||||          |||          :|||          |:!:
Db      67  SVNLN-DASNLDHDLASCYLFRNGSSEENSVIASAESKEAKLSRQR--TPEK-----116
         |||||          |||          :|||          |:!:
Qy     181  SDLESFLAGHTGPSLPAAATPAELIVPPAEIEICEFFFAAQAQAKRSKNPDFVRGV 240
         |||||          |||          :|||          |:!:
Db     117  -----|-----MPSEXIEIFFFAARQAILKRPRKKYNFDFEKEE 150
         |||||          |||          :|||          |:!:
Qy     241  PLDAGGRFEWAPVVVS 255
         ||:||||:|
Db    151  PLE--GREYEWIRIGS 163

```

RESULT 14
Q9LJLS
ID Q9LJLS PRELIMINARY; PRT; 196 AA.
AC AC
DT DT 01-OCT-2000 (T+EMBLrel. 15, Created)
DT DT 01-OCT-2000 (TREMBUrel. 15, Last sequence update)
DE DE 01-OCT-2003 (TREMBUrel. 25, Last annotation update)
DE DE GB AAC23758.1 (Cyclin-dependent kinase inhibitor 6).
OX KRP6 OR ACKI.
RN OS Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosoids II; Brassicales; Brassicaceae; Arabidopsi.
NCBI_TaxID=3702;
[1] RN
RN RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RC Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
[2] RN
RN RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RY MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAG and BAC clones";
RL DNA Res. 7:217-221(2000).
[3] RN
RN RP SEQUENCE FROM N.A.
RR MEDLINE=21342510; PubMed=11449057;
RX de Veylder L., Beeckman T., Beemster G.T.S., Krools L., Terras F.,
RA Landrieu I., Van Der Schueren E., Maes S., Naudts M., Inze D.;
RA "Functional analysis of Cyclin-dependent kinase inhibitors of
RT Arabidopsis.";
RL Plant Cell 13:1653-1668(2001).
[4] RN
RN RP SEQUENCE FROM N.A.
RA Park S.C., Cho J.W.;
RA "A novel cyclin dependent kinase inhibitor (ackl).";
RT Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AF000419; BAB02955.1; -
DR EMBL; AJ301557; CAC41620.1; -
DR EMBL; AF106705; AAF77612.2; -
DR GO; GO:0005634; Cnucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . . ; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; F:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
KW Pfam; PF02234; CDI; 1.
XW Cyclin; Kinase.
SQ SEQUENCE 196 AA; 21454 MW; AC543304CEB92C4 CRC64;

```

RESULT 15
Q8S9U7 PRELIMINARY; PRT; 204 AA.
AC Q8S9U7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to cyclin-dependent kinase inhibitor (Putative cyclin-
DE dependent kinase inhibitor).
GN B1064G04.1 OR B1144D11.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]_TaxID=39947;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone:B1064G04.1";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone:B1144D11.1";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003924; BAB85380.1; -.
DR EMBL; AP003443; BAC00670.1; -.
DR Gramene; Q8S9U7; -.
DR GO; GO:0016301; F:kinase activity; IEA.
KW Cyclin; Kinase.
SQ SEQUENCE 204 AA; 21233 MW; 900E2DDDA5299150 CRC64;

```

Query Match	9.9%;	Score 129.5;	DB 10;	Length 204;
Best Local Similarity	27.5%;	Pred. No. 0.17;		
Matches	61;	Conservative 18;	Mismatches 50;	Indels 93; Gaps 12;
QY	1	MGKYMRKCRGAAGA--EVAAEVTQVGVRTSRSAATGGVAKVAPRKRKAPAGEPAA	58	
DB	1	MGKYMRQAKVZAAVWELAAV-APAPLGVTRTRSLAL-----QKR-----	40	
QY	59	VSAGDGGSCYIHLRSMFLMAPPQPSVDSVP---TPVEAADGAAGQGAALAGLSR	115	
DB	41	---QGGE---YLELRSRRLKLPDPDPDP---PRRRAPVTTDPAPAA-----	78	
QY	116	CSSTASSVNLGLGGQGRGSHTRCSYDAAAGDGHVLVDVSAASNSGSGPDRRRTTBSR	175	
DB	79	-----ESAEAG-----SFGGGER-----PRAG	97	
QY	176	ANGELSDLESDLAGHKTFGL--PAATPAEELIV---PPAHE	212	
DB	98	GHGKVVQSLADQPGITGVLDPPFAAKHATGTVADRPLHK	139	

Mon Oct 4 08:49:19 2004

us-09-993-808b-2.rspt

Page 8

Search completed: October 1, 2004, 16:53:48
Job time : 119 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:38:10 ; Search time 121 Seconds
(without alignments)
597.786 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 1304

Sequence: 1 MGKYMRCRGAAGRAEVAE.....VRGVPLDAGGRFENAPVVS 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1304	100.0	256	6	ABB98757 Maize CKI
2	669	51.3	262	5	ABG65670 OsICK 2 p
3	665	51.0	227	6	ABB98759 Maize CKI
4	612	46.9	125	3	AAB26250 Corn cycl
5	411	31.5	126	3	AAB26249 Corn cycl
6	399.5	30.6	116	5	ABG65674
7	279	21.4	226	5	ABG65692 Rice OsIC
8	235	18.0	194	5	ABG65691 Rice OsIC
9	232.5	17.8	53	3	AAB26245 Corn cycl
10	232.5	17.8	53	3	AAP01940 Corn cycl
11	219	16.8	190	3	AAB27258 Wheat cyc
12	217	16.6	42	6	ABB98860 Maize CKI
13	215.5	16.5	222	5	AAE25111 Arabidops
14	204	15.6	190	3	AAB26252 Corn cycl
15	202.5	15.5	46	6	ABB98862 Maize CKI
16	200	15.3	190	6	ABB98758 Maize CKI
17	190.5	14.6	216	2	AAW98180 Arabidops
18	178	13.7	196	3	AAB27259 Chenopodi
19	178	13.7	196	3	AAE25108 Chenopodi
20	178	13.7	196	3	AAE25108 Chenopodi
21	178	13.7	289	3	AAB27262 Arabidops
22	178	13.7	289	3	AAG40485 Arabidops
23	178	13.7	289	5	AAE25109 Arabidops
24	178	13.7	289	5	AAU72581 Arabidops
25	160	12.3	205	3	AAP01952 Cyclin de

26	160	12.3	205	3	AAB27254
27	156	12.0	256	3	AAG40486 Arabidops
28	149	11.4	327	3	AAB27261 Arabidops
29	147.5	11.3	136	3	AAE25109 Arabidops
30	145.5	11.2	37	3	AAP01950 Rice cycl
31	145.5	11.2	37	3	AAB27252 Rice cycl
32	145	11.1	137	5	AAE25107 Arabidops
33	144.5	11.1	108	5	ABG65676 Broom cor
34	142.5	10.9	54	3	AAB26253 Corn cycl
35	137	10.5	201	3	AAG40487 Arabidops
36	134	10.3	90	5	ABG65673 Rice Os I
37	134	10.3	99	3	AAB27251 Rice cycl
38	134	10.3	99	3	AAP01949 Cyclin de
39	134	10.3	196	5	AAE25104 Arabidops
40	134	10.3	208	4	AAB47001 Plant D-1
41	134	10.3	212	5	AAE25103 Arabidops
42	134	10.3	212	5	AAE25103 Arabidops
43	129.5	9.9	93	5	ABG65675 Corn ICK
44	129	9.9	88	3	AAB27257 Soybean c
45	126	9.7	60	3	AAP01953 Cyclin de

ALIGNMENTS

RESULT 1

ABB98757
ID ABB98757 standard; protein; 256 AA.

XX AC ABB98757;

XX 21-FEB-2003 (first entry)

XX DE Maize CKI_B.

XX KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI B; plant; corn;
crop yield; root size; plant growth; tassel size; ear size;
male sterility; endoreduplication.

XX OS Zea mays.

XX FH Key Location/Qualifiers

FT Domain 1..7
FT /note= "Conserved domain"

XX EN WO200281623-A2.

XX PD 17-OCT-2002.

XX PF 06-NOV-2001; 2001WO-US044038.

XX PR 07-NOV-2000; 2000US-0246349P.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

XX DR WPI; 2003-058511/05.

XX DR N-PSDB; ABV74603.

XX PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
involved in cell cycle regulation, and useful for altering cell cycle
protein content, cell cycle progression, cell number and composition of
plants.

XX PS Claim 12; Page 65-66; 69pp; English.

XX CC The present sequence is maize cyclin-dependent kinase inhibitor (CKI),
CKI_B. The coding sequence for this protein (I) is useful for modulating
the activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
barley, oil-seed Brassica and millet. Modulating the activity of CDK,

CC preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain
 XX

SQ Sequence 256 AA;
 Query Match 100.0%; Score 1304; DB 6; Length 256;
 Best Local Similarity 100.0%; Pred. No. 2.3e-109; Indels 0; Gaps 0;
 Matches 256; Conservative 0; Mismatches 0;
 QY 1 MGKMKRCKGAGAEVAEVTQVVGRTSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
 DB 1 MGKMKRCKGAGAEVAEVTQVVGRTSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
 QY 61 AGDGGSCYIHLRSRMLFWAPPQPSVDSVPTPVEADGAGCQQAALAGLSRCSSTA 120
 DB 61 AGDGGSCYIHLRSRMLFWAPPQPSVDSVPTPVEADGAGCQQAALAGLSRCSSTA 120
 QY 121 SSVNLGLGGRGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRERTTSSRAHGE 180
 DB 121 SSVNLGLGGRGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRERTTSSRAHGE 180
 QY 181 SDLESGLAGHKTGPSLPAATPAELIIVPPAHEIQEFAAEAAQAKRFASKYNFDFVRGV 240
 DB 181 SDLESGLAGHKTGPSLPAATPAELIIVPPAHEIQEFAAEAAQAKRFASKYNFDFVRGV 240
 QY 241 PLDAGGRFEPVWSI 256
 DB 241 PLDAGGRFEPVWSI 256

RESULT 2
 ABG65670
 ID ABG65670 standard; protein; 262 AA.

XX AC ABG65670;
 XX DT 27-AUG-2002 (first entry)
 XX OSICK 2 protein.
 XX KW Plant; inhibitor of cyclin dependent kinase; ICK.
 XX OS Oryza sativa.
 XX XX WO200228893-A2.
 XX PD 11-APR-2002.
 XX PF 29-JUN-2001; 2001WO-IB001492.
 XX PR 14-JUL-2000; 2000US-0218471P.
 XX PR 13-OCT-2000; 2000US-0241219P.
 XX PA (CROP-) CROPDISEGN NV.
 XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 XX PI Hatzfeld Y;
 XX DR WPI; 2002-471311/50.
 XX DR N-PSDB; ABK93958.

XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.
 XX Claim 48; Fig 3; 141pp; English.
 XX This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents an inhibitor of
 CC cyclin dependent kinase (ICK) protein of the invention
 XX

SQ Sequence 262 AA;
 Query Match 51.3%; Score 669; DB 5; Length 262;
 Best Local Similarity 60.3%; Pred. No. 4.9e-52;
 Matches 161; Conservative 19; Mismatches 67; Indels 20; Gaps 10;
 QY 1 MGKMKRCKGAGAEVAEVTQVVGRTSRSAATGG-----VAKVAPRRKRAPAGE 54
 DB 1 MGKMKRCKGAGAEVAEVTQVVGRTSRSAATGGATTTKVQAAASASTRRRKALLP 60
 QY 55 PAAAVSAGDGGSCYIHLRSRMLFWAPPQPSVDSVPTPVEADGAGCQQAALAGLS 114
 DB 61 TAVVGTTRDGGSCYIHLRSRMLFWAPPQPSVDSVPTPVEADGAGCQQAALAGLS 118
 QY 115 RCSSTASSVNLGLGGRGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRERTTSS 174
 DB 119 RCSSTASSVD--AAADRSLACRS-DVAEAGEH--VPEGSASDSASGRDRERTTSS 173
 QY 175 RAHGELSDLESILA-GHKTGPSILP-AATPAELI-----VPPAHEIQEFAAEAAQAKRF 228
 DB 174 FLPGVEVSDLESILAGGQKSRPLPSAATASQAQATRPKIPPAAEIEFAFAAEAEAKRF 233
 QY 229 ASKYNFDFVRGVPLDAGGRFEPVWSI 255
 DB 234 AAKYNFDFVRGVPLDAGGRFEPVWSI 259

RESULT 3
 ABB98759
 ID ABB98759 standard; protein; 227 AA.

XX AC ABB98759;
 XX DT 21-FEB-2003 (first entry)
 XX XX Maize CKI_D.
 XX KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_D; plant; corn;
 XX KW crop yield; root size; plant growth; tassel size; ear size;
 XX KW male sterility; endoreduplication.
 XX OS Zea mays.
 XX FH Key Location/Qualifiers
 XX FT Domain 1..7
 XX FT /note= "Conserved domain"
 XX FT Misc-difference 111
 XX FT /note= "Encoded by CAR"
 XX XX WO200281623-A2.
 XX XX 17-OCT-2002.
 XX PF 06-NOV-2001; 2001WO-US044038.

```
XX 07-NOV-2000; 2000US-0246349P.
XX (PION-) PIONEER HI-BRED INT INC.
XX (ARIZ-) ARIZONA BOARD OF REGENTS.
XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
XX N-PSDB; ABV74605.
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
XX involved in cell cycle regulation, and useful for altering cell cycle
XX protein content, cell cycle progression, cell number and composition of
XX plants.
XX Claim 12; Page 69; 69pp; English.
XX The present sequence is maize cyclin-dependent kinase inhibitor (CKI),
XX CKI D. The coding sequence for this protein (I) is useful for modulating
XX the activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
XX soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
XX barley, oil-seed Brassica and millet. Modulating the activity of CDK,
XX preferably modulating downward is useful for providing differential
XX growth in a plant, especially a positive growth advantage and modulating
XX CDK activity upward is useful for increasing crop yield, root size, plant
XX growth, tassel size and/or ear size. Modulating CDK activity is also
XX useful for conferring male sterility and for improving transformation
XX frequencies by increasing the number of cells in cell division. CDK
XX activity can also be modulated for modulating endoreduplication in the
XX endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
XX promoter used is an endosperm-preferred promoter. The cell numbers are
XX modulated in one or more tissues of a plant, comprising root, seed,
XX tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
XX stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
XX pericarp, meristematic or leaf cells. (I) is also useful for identifying
XX maize CKI interacting proteins, by adducting the nucleic acid sequence to
XX a second nucleic acid sequence encoding a DNA-binding domain
XX Sequence 227 AA;
XX
XX Query Match 51.0%; Score 665; DB 6; Length 227;
XX Best Local Similarity 72.6%; Pred. No. 9.3e-52;
XX Matches 172; Conservative 6; Mismatches 31; Indels 28; Gaps 16;
XX
XX QY 1 MGKTYMKRCRGAGAEVAAVEVTVGVVTRSRSAAT-GGVAKVA-PRRKAA--PA---- 52
XX Db 1 MGKTYMKRERGAGGAEVAAVEVTVGVVTRSRSAATGGGVAKVAPRRKXKALLPAANVT 60
XX
XX QY 53 --GEPAAVAGAGGGGS-CYTHLSRMLFMAPPQPSVDSVPTPVEAADCAAGQGG-AA 108
XX Db 61 TSGEFGAVAGGGGGGGGCCYHLSRMLFMAPPQGPS--AALTPTVEAA-GAAKQGGVVA 117
XX
XX QY 109 LAAGLSRCSSTASSVNLGLGQGRSHTCRSYDA-AEAGGDHVLVDVSAASNSGSGPDRER 167
XX Db 118 LAAGLSRCSSTASSVDV-----GGHACRSDAAPAEVDGDHV-PDVTASNSGSGVPDRER 170
XX
XX QY 168 RETTP-SSRAH-GLSLDLSLAG-HKYG-PSLPA-ATPAELIIVPPAHEIQEFPFAA 219
XX Db 171 RETTPSSSRAHGGELSLDLSLVGRQKTCGSSSPATTTSAELIIVPPAHEIQEFPFAA 227
XX
XX RESULT 4
XX AAB26250
XX ID AAB26250 standard; protein; 125 AA.
XX AC AAB26250;
XX XX
XX 17-JAN-2001 (first entry)
XX DE Corn cyclin-dependent kinase inhibitor #3.
XX KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
```

```
XX cell growth; herbicide.
XX OS Zea mays.
XX PN WO200060087-A2.
XX PD 12-OCT-2000.
XX PF 06-APR-2000; 2000WO-US009106.
XX PR 07-APR-1999; 99US-0128192P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI Klein TM, Weng Z, Cahoon RE;
XX WPI; 2000-679375/66.
XX N-PSDB; AAA95281.
XX Cyclin dependent kinase inhibitor sequences, useful for identifying
XX herbicides and plant growth inhibitors.
XX PS Claim 10; Fig 1; 58pp; English.
XX The present sequence is the corn cyclin-dependent kinase inhibitor
XX (CDKI). Its coding sequence was isolated by searching a contig comprising
XX cDNA from corn roots and ear leaf sheath for sequences similar to those
XX encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and
XX Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote
XX or inhibit cell division and growth. The protein and its coding sequence
XX are useful in the production of transgenic plants which produce increased
XX or decreased amounts of the CDKI protein, in the identification of
XX herbicides, in genetic and physical mapping and in the isolation of the
XX CDKI gene in other organisms
XX SQ Sequence 125 AA;
XX
XX Query Match 46.9%; Score 612; DB 3; Length 125;
XX Best Local Similarity 96.0%; Pred. No. 2.6e-47;
XX Matches 120; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 132 GSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHGLSLDLSLAGHK 191
XX Db 1 GSHTCRSDDAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHGLSLDLSLAGHK 60
XX
XX QY 192 TGPLPAATPAAEILIVPPAHEIQEFPFAAEEAAQAKRFASKYNFDFVRGVPLDAGRFEWA 251
XX Db 61 TGPLPAATPAAEILIVPPAHEIQEFPFAAEEAAQAKRFASKYNFDFVRGVPLDAGRFEWA 120
XX
XX QY 252 PWVSI 256
XX Db 121 PWVSI 125
XX
XX RESULT 5
XX AAB26249
XX ID AAB26249 standard; protein; 126 AA.
XX AC AAB26249;
XX XX
XX 17-JAN-2001 (first entry)
XX DE Corn cyclin-dependent kinase inhibitor #2.
XX KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
XX cell growth; herbicide.
XX OS Zea mays.
XX PN WO200060087-A2.
XX PD 12-OCT-2000.
XX
```

PF 06-APR-2000; 2000WO-US009106.
 XX
 PR 07-APR-1999; 99US-0128192P.
 XX
 PA (DUPO) DU PONT DE NEWMOURS & CO E I.
 XX
 PI Klein TM, Weng Z, Cahoon RE;
 XX
 XX WPI; 2000-679375/66.
 DR N-PSDB; AAA95280.
 XX
 XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.
 XX
 XX Claim 10; Fig 1; 58pp; English.
 XX
 CC The present sequence is the corn cyclin-dependent kinase inhibitor
 CC (CDKI). Its coding sequence was isolated by searching a corn silk cDNA
 CC library for sequences similar to those encoding the CDKI from Chenopodium
 CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
 CC in the cell cycle, and may promote or inhibit cell division and growth.
 CC The protein and its coding sequence are useful in the production of the
 CC transgenic plants which produce increased or decreased amounts of the
 CC CDKI protein, in the identification of herbicides, in genetic and
 CC physical mapping and in the isolation of the CDKI gene in other organisms
 XX
 XX Sequence 126 AA;
 SQ
 Query Match 31.5%; Score 411; DB 3; Length 126;
 Best Local Similarity 78.7%; Pred. No. 3.6e-29;
 Matches 100; Conservative 1; Mismatches 18; Indels 8; Gaps 8;
 QY 136 CRSDYDA-ARAGDHLVDVSAASNSGSGPDRERETTP-SSRAH-GEISDLESDLAG-HK 191
 DB 2 CRSDAAPAEVDGHDV-PDVVTASNSGSGVPDRERETTPSSSRAHGGELSDLESDLVGROK 60
 QY 192 TG-PSLPAATPA-AELIVPPAHEIQEFFFFAAAHAAQAKRFASKYNFDFVRGVPLDAGRPE 249
 DB 61 TCSSSPATTTVAELIVPPAHEIQEFFFFAAAHAAQAKRFASKYNFDFVRGVPLDA-GRPE 119
 QY 250 WAPVSVI 256
 DB 120 WTGVSII 126
 RESULT 6
 ABG65674
 ID ABG65674 standard; protein; 116 AA.
 AC ABG65674;
 XX
 XX 07-AUG-2003 (revised)
 DT 27-AUG-2002 (first entry)
 XX
 XX Corn ICK 1 protein.
 XX
 XX Plant; inhibitor of cyclin dependent kinase; ICK.
 XX
 XX Zea mays.
 OS
 XX WO200228893-A2.
 PN
 XX 11-APR-2002.
 PD
 XX 29-JUN-2001; 2001WO-IB001492.
 PF
 XX 14-JUL-2000; 2000US-0218471P.
 PR 13-OCT-2000; 2000US-0241219P.
 XX
 XX (CROP-) CROPDESIGN NV.
 PA
 XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 PI Hatzfeld Y;

XX WPI; 2002-471311/50.
 DR N-PSDB; ABK93954.
 XX
 PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.
 XX
 XX Claim 48; Fig 4; 141pp; English.
 XX
 CC This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents an inhibitor of
 CC cyclin-dependent kinase (ICK) protein of the invention. (Updated on 07-
 CC AUG-2003 to correct OS field.)
 XX
 XX Sequence 116 AA;
 SQ
 Query Match 30.6%; Score 399.5; DB 5; Length 116;
 Best Local Similarity 81.9%; Pred. No. 3.6e-28;
 Matches 95; Conservative 0; Mismatches 14; Indels 7; Gaps 7;
 QY 146 GDHVLVDVSAASNSGSGPDRERETTP-SSRAH-GEISDLESDLAG-HKTG-PSLPA-AT 200
 DB 3 GDHV-PDVVXASNSGSGVPDRERETTPSSSRAHGGELSDLESDLVGROKTCSSSPATTT 61
 QY 201 PRAELIVPPAHEIQEFFFFAAAHAAQAKRFASKYNFDFVRGVPLDAGRPEWAPVSI 256
 DB 62 SAELIVPPAHEIQEFFFFAAAHAAQAKRFASKYNFDFVRGVPLDA-GRFEWTPGVSI 116
 RESULT 7
 ABG65692
 ID ABG65692 standard; protein; 226 AA.
 XX
 XX AC ABG65692;
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX Rice OsICK5 protein.
 DE
 XX Plant; inhibitor of cyclin dependent kinase; ICK.
 KW
 XX Oryza sativa.
 OS
 XX WO200228893-A2.
 PN
 XX 11-APR-2002.
 PD
 XX 29-JUN-2001; 2001WO-IB001492.
 PF
 XX 14-JUL-2000; 2000US-0218471P.
 PR 13-OCT-2000; 2000US-0241219P.
 XX
 XX (CROP-) CROPDESIGN NV.
 PA
 XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 PI Hatzfeld Y;
 XX
 XX WPI; 2002-471311/50.
 DR N-PSDB; ABK93981.
 XX
 XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 XX to screen substrates, drugs or compounds which modulate ICK activity and
 XX treat disorders characterized by an insufficient or excessive production
 XX of ICK inhibitors.
 PT

XX PS Claim 48; Page 139-140; 141pp; English.

XX CC This invention relates to the DNA and protein sequences of novel isolated

CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of

CC the invention may be used for treating disorders characterised by

CC insufficient or excessive production of an ICK inhibitor. The protein of

CC the invention may also be used to screen for naturally-occurring ICK

CC substrates, drugs or compounds which modulate ICK activity, as well as to

CC treat disorders characterised by insufficient or excessive production of

CC ICK protein, forms which have decreased or aberrant activity compared to

CC ICK wild type protein. The present sequence represents an inhibitor of

CC cyclin dependent kinase (ICK) protein of the invention

XX SQ Sequence 226 AA;

Query Match 21.4%; Score 279; DB 5; Length 226;

Best Local Similarity 34.4%; Pred. No. 6.4e-17;

Matches 98; Conservative 26; Mismatches 71; Indels 90; Gaps 13;

QY 1 MGKYMRCRGAAGAEVAEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAVS 60

Db 1 MGK-KKKRDGAARQARVVG---GVTR-----AAVTARRVVASAEGCGLVG 46

QY 61 AGDGGG-----CYIHLRSRMFLFMAAPPQPSVDSVPTVEAADGAAGQGAALAG 112

Db 47 RGGGGGGDDGGGCGYLRLASRL-----PFVAAAVVSSRREALGDS 90

QY 113 LSRCSSTASSVNLGLGGGSHTCRSYDAAGAGDHVLVDVSAASNSGSGPDRERTTP 172

Db 91 VAESAASSSSRAVELLG---C-----SGEEAAMAEKXSA-----TTP 124

QY 173 SSR---AHGELSDLES-----LAGHKTGPSLPAATPAELIVPP---AHE 212

Db 125 SSRPPGADSDSAESNOEAQOMCRSSSTTSAAAFHAG-----ATTRFRMWAPPAAAE 180

QY 213 IQEFPFAAAEAQAFASKYNFDFVRGVPLDAG--GRFEWAPVVS 255

Db 181 IEFEFLAAERSEARFAAKYNFVVRGVPLDAGGAGFEWTA VGS 225

RESULT 8

ABG65691

ID ABG65691 standard; protein; 194 AA.

AC ABG65691;

DT 27-AUG-2002 (first entry)

XX Rice OsICK4 protein.

DE Plant; inhibitor of cyclin dependent kinase; ICK.

XX Oryza sativa.

XX WO200228893-A2.

PN 11-APR-2002.

XX 29-JUN-2001; 2001WO-IB001492.

XX 14-JUL-2000; 2000US-0218471P.

PR 13-OCT-2000; 2000US-0241219P.

XX (CROP-) CROPDDESIGN NV.

PA Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;

PI Hatzfeld Y;

XX WPI; 2002-471311/50.

DR N-PSDB; ABK93971.

XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used

PT to screen substrates, drugs or compounds which modulate ICK activity and

PT treat disorders characterized by an insufficient or excessive production

PT of ICK inhibitors.

XX Claim 48; Fig 3; 141pp; English.

XX This invention relates to the DNA and protein sequences of novel isolated

CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of

CC the invention may be used for treating disorders characterised by

CC insufficient or excessive production of an ICK inhibitor. The protein of

CC the invention may also be used to screen for naturally-occurring ICK

CC substrates, drugs or compounds which modulate ICK activity, as well as to

CC treat disorders characterised by insufficient or excessive production of

CC ICK protein, forms which have decreased or aberrant activity compared to

CC ICK wild type protein. The present sequence represents an inhibitor of

CC cyclin dependent kinase (ICK) protein of the invention

XX SQ Sequence 194 AA;

Query Match 18.0%; Score 235; DB 5; Length 194;

Best Local Similarity 31.5%; Pred. No. 4.9e-13;

Matches 82; Conservative 25; Mismatches 73; Indels 80; Gaps 12;

QY 1 MGKYMRCRGAAGAEVAEVTQV---GVTRSRSAATGGVAKVAPRRKRAPAGEPA 56

Db 1 MGKYMRAKVVVSSEVVAAYWELAAAPLGVTRARSLAL-----QKR----- 43

QY 57 AAVSAGDGGSCYIHLRSRMFLFMAAPPQPSVDSVPTVEAADGAAGQGAALAGLSRC 116

Db 44 ---QGGE---YLELSRRLEKLPPPP-----PPPRRATAAA-----ATA 78

QY 117 SSTASSVNLGLGGGSHTCRSYDAAEA---GGDHVLVDVSAASNSGSGPDRERTTP 172

Db 79 DATAT-----ESAEAVSFGGENVL-ELEAM-----ERNTRTTP 112

QY 173 SS---RAHGELSDLESDLAGHKTGPSLPAATPAELIVPPAHEIOEFFFAAAEAQAFAS 230

Db 113 CSLIRDPTTSTPGSTTTRSHSSHCKVQTPVRNIIIPASALEFAFFAAEQRQAFID 172

QY 231 KYNFDFVRGVPLDAGGRFEW 250

Db 173 KYNFDPVNDCLP--GRFEW 190

RESULT 9

AAB26245

ID AAB26245 standard; protein; 53 AA.

AC AAB26245;

DT 17-JAN-2001 (first entry)

XX Corn cyclin-dependent kinase inhibitor #1.

DE Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;

XX cell growth; herbicide.

XX Zea mays.

XX WO2000060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU FONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX N-PSDB; AAA95276.

DR


```

Db 46 AADGALPNDSCYQLRRERLEKPSLLIEPKQP-PRVH-----RSIGKESG 90
QY 112 GLSRCSSTASSVNLGLGQGRSHTCRSYDAABAGDHLVDVSAASNS- GSGPDRERRET 170
Db 91 SRSRVDVSNS-----APVAQSNEDFCDFNVSVQVSCGENSLGFESRHSSTRES 139
QY 171 TPSSRAHGLSDLESDLAGHKTGPSLPAATP-----AABELIVPPAHEIOEFFFAAAAEAAQAK 226
Db 140 TPCN----FVDEMEINVTGSTRSMCRATKEYTREQDNVPIPTSEMEEFFAYAEQQOOR 195
QY 227 RFASKYNFDFVRGVPLDAGGRPEWAPV 253
Db 196 LFMEXYNFIDVNDIPL--SGRIEYVQV 220

RESULT 14
ID AAB26252 standard; protein; 190 AA.
XX AC AAB26252;
XX DT 17-JAN-2001 (first entry)
XX DE Corn cyclin-dependent kinase inhibitor #5.
XX KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
XX KW cell growth; herbicide.
XX OS Zea mays.
XX PN WO200060087-A2.
XX PD 12-OCT-2000.
XX PF 06-APR-2000; 2000WO-US009106.
XX PR 07-APR-1999; 99US-0128192P.
XX PA (DUPO ) DU PONT DE NEWMOURS & CO E I.
XX PI Klein TM, Weng Z, Cahoon RE;
XX WPI; 2000-679375/66.
XX N-PSDB; AAA95283.

Cyclin dependent kinase inhibitor sequences, useful for identifying
herbicides and plant growth inhibitors.
Claim 10; Fig 1; 58pp; English.

The present sequence is the corn cyclin-dependent kinase inhibitor
(CDKI). Its coding sequence was isolated by searching a contig of corn
endosperm, leaf sheath, tassel and shoot cDNA for sequences similar to
those encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans
and Arabidopsis thaliana. CDKI is involved in the cell cycle, and may
promote or inhibit cell division and growth. The protein and its coding
sequence are useful in the production of transgenic plants which produce
increased or decreased amounts of the CDKI protein, in the identification
of herbicides, in genetic and physical mapping and in the isolation of
the CDKI gene in other organisms.

Sequence 190 AA;

Query Match 15.6%; Score 204; DB 3; Length 190;
Best Local Similarity 28.2%; Pred. No. 3e-10;
Matches 71; Conservative 34; Mismatches 79; Indels 68; Gaps 10;

QY 1 MGKYMKRCGAAGAEVAABVTQV-VGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAV 59
Db 1 MGKYMRAK--ASSEVVIMDVAAAPLGVTRRALAL-----QLQEQQTQWER 47
QY 60 SAGDGGSCYIHLRSRMLFNAPPQPSVDSVPTFVEADGAGAQOQGAALAGLSRCGST 119

```

```

Db 48 GAGGE---YLELRNRLEKLPPLP-----ATT 71
QY 120 ASSVNLGLGQGRSHTCRSYDAABAGDHLVDVSAASNSGSGPDRERRETTPSSRAHGE 179
Db 72 RRS-----GGRKAAAEAAATKEAEASYGENMLEAM-----ERITRETTPCSLINTQ 119
QY 180 LSDLE-SDLAGHKTGPSLPAATPAAELIVPPAHEIOEFFFAAAAEAAQAKREASKYNFDFVR 238
Db 120 MSTPGSTFSGHSCHRRVNAFPVHA--VPSRENNFYFAAFQRQOQDFIDKYNFDPAN 176
QY 239 GVPLDAGGRFEW 250
Db 177 DCPLP--GRFEW 186

RESULT 15
ID ABB98862 standard; protein; 46 AA.
XX AC ABB98862;
XX DT 21-FEB-2003 (first entry)
XX DE Maize CKI_D C-terminal protein fragment.
XX KW Maize; cyclin-dependent kinase inhibitor; CKI; plant; corn; crop yield;
XX KW root size; plant growth; tassel size; ear size; male sterility;
XX KW endoreduplication.
XX OS Zea mays.
XX PN WO200281623-A2.
XX PD 17-OCT-2002.
XX PF 06-NOV-2001; 2001WO-US044038.
XX PR 07-NOV-2000; 2000US-0246349P.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.

Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
involved in cell cycle regulation, and useful for altering cell cycle
protein content, cell cycle progression, cell number and composition of
plants.

Example 11; Page 46; 69pp; English.

The present invention relates to maize cyclin-dependent kinase inhibitor
(CDKI), CKI_B, CKI_C and CKI_D. The coding sequences (I) for these
proteins are useful for modulating the activity of CDK in a plant such as
a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton,
rice, barley, oil-seed Brassica and millet. Modulating the activity of
CDK, preferably modulating downward is useful for providing differential
growth in a plant, especially a positive growth advantage and modulating
CDK activity upward is useful for increasing crop yield, root size, plant
growth, tassel size and/or ear size. Modulating CDK activity is also
useful for conferring male sterility and for improving transformation
frequencies by increasing the number of cells in cell division. CDK
activity can also be modulated for modulating endoreduplication in the
endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
promoter used is an endosperm-preferred promoter. The cell numbers are
modulated in one or more tissues of a plant, comprising root, seed,
tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
pricarip, meristematic or leaf cells. (I) is also useful for identifying
maize CKI interacting proteins, by adducting the nucleic acid sequence to

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:49:11 ; Search time 32 Seconds
(without alignments)
413.008 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVELDAGGEFAPVWSI 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCRTUS COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	9.5	826	4	US-09-894-998A-47
2	117.5	9.0	456	4	US-09-252-991A-19417
3	112	8.6	477	4	US-09-252-991A-24574
4	111	8.5	317	4	US-09-252-991A-30984
5	111	8.5	425	4	US-09-252-991A-31183
6	111	8.5	455	4	US-09-252-991A-18699
7	110.5	8.5	1239	4	US-09-252-991A-30198
8	109.5	8.4	605	4	US-09-252-991A-27220
9	109	8.4	300	4	US-09-252-991A-25073
10	108.5	8.3	726	4	US-09-252-991A-20675
11	108.5	8.3	973	4	US-09-252-991A-21386
12	108.5	8.3	1395	4	US-09-252-991A-30345
13	108	8.3	362	4	US-09-252-991A-17884
14	108	8.3	553	4	US-09-252-991A-32621
15	107.5	8.2	534	4	US-09-252-991A-32637
16	107.5	8.2	698	3	US-08-941-445A-11
17	107	8.2	433	4	US-09-252-991A-20044
18	106	8.1	560	4	US-09-252-991A-27797
19	106	8.1	743	4	US-09-252-991A-28327
20	106	8.1	871	4	US-09-252-991A-19431
21	106	8.1	1046	4	US-09-252-991A-16902
22	105.5	8.1	646	4	US-09-252-991A-28206
23	104.5	8.0	478	4	US-09-252-991A-19635
24	104.5	8.0	909	3	US-09-425-383-2
25	104.5	8.0	2294	4	US-09-252-991A-17231
26	104	8.0	395	4	US-09-252-991A-17341
27	103	7.9	316	4	US-09-252-991A-30877

28	103	7.9	681	4	US-09-430-291-4	Sequence 4, Appli
29	103	7.9	686	4	US-09-252-991A-20509	Sequence 20509, A
30	103	7.9	691	4	US-09-430-291-6	Sequence 6, Appli
31	102.5	7.9	445	4	US-09-252-991A-23505	Sequence 23505, A
32	102.5	7.9	553	3	US-09-083-351-2	Sequence 2, Appli
33	102.5	7.9	553	3	US-09-083-352-2	Sequence 2, Appli
34	101	7.7	196	4	US-09-252-991A-23071	Sequence 23071, A
35	101	7.7	214	4	US-09-252-991A-21304	Sequence 21304, A
36	101	7.7	297	4	US-09-252-991A-27909	Sequence 27909, A
37	101	7.7	335	2	US-08-405-175A-6	Sequence 6, Appli
38	101	7.7	618	4	US-09-252-991A-27666	Sequence 27666, A
39	100.5	7.7	347	4	US-09-252-991A-19498	Sequence 19498, A
40	100.5	7.7	738	3	US-08-864-038A-3	Sequence 3, Appli
41	100.5	7.7	931	4	US-09-252-991A-22550	Sequence 22550, A
42	100.5	7.7	1093	3	US-08-545-860D-55	Sequence 55, Appl
43	100.5	7.7	1093	5	PCT-US94-04496-55	Sequence 55, Appl
44	100	7.7	311	4	US-09-252-991A-18324	Sequence 18324, A
45	100	7.7	353	4	US-09-252-991A-32769	Sequence 32769, A

ALIGNMENTS

RESULT 1
US-09-894-998A-47
; Sequence 47, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HSV-2
US-09-894-998A-47

Query Match	9.5%	Score 124;	DB 4;	Length 826;
Best Local Similarity	28.3%	Pred. No. 0.002;		
Matches	60;	Conservative	18;	Mismatches 82; Indels 52; Gaps 9;
QY	6	RKCRGAGAEVAEVTQVGVTRSRSAATGGVAKVAPRRKEAPAGE-----PAAA 58		
Db	514	RKRGDSGSPAASSASS--SAAPRSLAPQGVGAKAAAPR--RAPDSGDRGHGLAP 569		
QY	59	VSAGGGGSCYIHLRSMLFMAPPOPSVDSVPTPEAADGAAGQGAALACLRCSS 118		
Db	570	ASAGA-----APPSASPS-----SQAAVAASSSSASSSSASSSSAS 606		
QY	119	TASVNLGLGGQSGHTRCSYDAEAGGDHVLVDVSAANSNGSGPDRETTPTSSRAHG 178		
Db	607	SSASSSSSASSSSASSSSASSAGGAG-----SVASASGAG--ERRETSIGPAAA 656		
QY	179	ELSDLESDLAGHKT-----GPSLPATPAEL 205		
Db	657	PRGPRK---CARKTRAEAGGFEPCARDPAFGL 685		

RESULT 2
US-09-252-991A-19417
; Sequence 19417, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19417
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (62)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-19417

Query Match 9.0%; Score 117.5; DB 4; Length 456;
Best Local Similarity 29.9%; Pred. No. 0.0038;
Matches 66; Conservative 13; Mismatches 73; Indels 69; Gaps 13;
QY 26 GVRTSRGAATGGVAKVAPRRKAPA-----GEP-----AAVSAGDGGSCVILHR 73
DB 233 GERRGAPASAGGAA--GKRCRPPARRQGTETEPATGCGGAPAPGGAAGS-----R 305
QY 74 SRMLFMAP-----PQPQSVDSVPTTVEAADGAGGCGAALAGL-----113
DB 306 GGATGAPRPGRYNRPRTIRRRPAQAPAPVRHA-GAAAAAGGGAAGLGGQLQGGDP 364
QY 114 -----SRCSSTASSVNLGLG-----ORGSHTCRSYDAAEAGDGHVLDVSAAS 157
DB 365 ARTHPSRATGRAG-----GGHVRPLAPRRQRGAGLRRGIRAAAGQRHLQHDRPA 418
QY 158 NSGS-GPDRETTTPSSRAHGLSDLESDLAGH-KTGPSL 196
DB 419 LPGAYPPGRRRRRAGPA--ASGDVGDH-----RGHSRTGSL 454

RESULT 3
US-09-252-991A-24574
; Sequence 24574, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24574
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24574

Query Match 8.6%; Score 112; DB 4; Length 477;
Best Local Similarity 25.5%; Pred. No. 0.014;
Matches 70; Conservative 22; Mismatches 91; Indels 92; Gaps 14;
QY 4 YMRCKRGA-----AGAEVAAVEVTQVGVTRSRSAATGGVAKVAPRRK-----48
DB 31 WWRACRSARWSPATSTTGASALNCSPVAACAKSSAPRAGRRAASRRGCCWTF 90
QY 49 --RAPAGEPAAVSGDGGSCVILHR-----SRMLFMAPQPQSVDSVPTTVP 95

DB 91 TCAMPAGASRAS-SRGGPGRCTCTTWRWRWSHELPLARRQPRRTADQRRRVLSAPV 149
QY 96 EAADGAAGQ-----QGAALAG-----LSRCSSTASSVNL-----GLGGG 132
DB 150 HRR-GAPGDPPGNLHHLRGRPAVAGSPQRRRRARCRGAGHGGRLHRLAQPGLP 208
QY 133 -----SHTCRSYDAAEAGDGHVLDVSAASNSGSGPDRE---RRETT 171
DB 209 RLGRSAPVRPAPAPARHAHQPVSPVAPQAGD-----RSPAGLRRRHQLRRG 258
QY 172 PSSRAHGLSDLESDLAGHKTGPS---LPAATPAA 203
DB 259 GTSRQHG-----QAGLRGARRPGGARYPPGLPGA 288

RESULT 4
US-09-252-991A-30984
; Sequence 30984, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30984
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30984

Query Match 8.5%; Score 111; DB 4; Length 317;
Best Local Similarity 25.3%; Pred. No. 0.0098;
Matches 73; Conservative 24; Mismatches 90; Indels 102; Gaps 13;
QY 13 GAEEVAAVEVTQVGVTR-----SRSAATGG--VAKVAPRRKRAPAGRP---AAAV 59
DB 4 GAAMPQMAHTAVPRARYRVHARRYLLSRGSGAGRWPAADVADRSAAADPGLEPLRAARR 63
QY 60 SAGDGGSCVILHRSMFLFMAPQPQSVDSVPTTV-----EAADGAAGQ 105
DB 64 RAGGGAAC---LQRRPGVPAPHPQAGAAVYAVGRRLSRHGFDPGLLEAGEGGAGHR 120
QY 106 GA-----ALAA-----GLSR-----CSSTASSVNLGLGQGRS---133
DB 121 PAQGRDPFELVATAAHRDADALRRRRPAPRFPGLPGCAERGETRLAAGRRRTGRQP 180
QY 134 -----HTCRSYDAAEAGDGHVLDVSAASNSGSGPDRERTTPSSRAHGLSDLES 189
DB 181 LPRERSARGLPQAG-----LAAASAATGAPRTTL-----LAR 215
QY 190 HKTGPSLPAATPAEELIVP-----PAHEIOE-----FFAAAEAAQAKR 227
DB 216 GAGTFAQPAAPRAARTHPVAAGAPAAAEQDPPGDRGHASAPRR 264

RESULT 5
US-09-252-991A-31183
; Sequence 31183, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 31183
;; LENGTH: 425
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31183

Query Match 8.5%; Score 111; DB 4; Length 425;
Best Local Similarity 23.2%; Pred. No. 0.015;
Matches 76; Conservative 24; Mismatches 104; Indels 124; Gaps 14;

QY 9 RGAGAEEV-AAVEVTVGVVTRSRSAATGKAVAPRRKRAPAGEPAAAVSAGDGG- 66
Db 106 RCAAGAAAGATAPPAAGLGRSQPVA---GAEEA-----QLFGGGFVAAVDFGRRGP 157
QY 67 -----SCVIHLRSRLFWA-----BPQOPSVDSV----- 91
Db 158 ARPVGQRRLATACLHRA-----MADRRDDPRRAGDPVLPHPRSRPSAAGDRLAGQ 213
QY 92 --PTPVAADGAAGQQAALAGLSRC-----SSTASSVNL----- 125
Db 214 WHPAGLRAG-GAHRGRHDAADALACKTVRRPRWCATVAPVALARSHLAGAALYSAP 272
QY 126 -----GLGGGSGHTCSDYDAAGAGGHVLDVYSAANSNGSGDPRE 166
Db 273 GIFIDRGAAMHGALEPQPGAGAEFGA-----ARAGGGRHGRVDPGDAGTGHRSFYRS 325
QY 167 RRETPSSRAHGELSDLESLLAGHKT-----GPSLPAATPAEELIVPPAHEIQEFAAA 220
Db 326 RAATA-----GSDALGFRSAQSLRGAGVPAQPVAGELGAAPGRRTVGVGVA 373
QY 221 E-----AAQAKRFASKYNFDFVRGVPLDA 244
Db 374 VRLVLCADAVSSAGRRPSRLIRGLAADA 401

RESULT 6
US-09-252-991A-18699
;; Sequence 18699, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; PRIOR FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 18699
;; LENGTH: 455
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18699

Query Match 8.5%; Score 111; DB 4; Length 455;
Best Local Similarity 24.9%; Pred. No. 0.016;
Matches 70; Conservative 26; Mismatches 127; Indels 58; Gaps 11;
QY 8 CRGAAGAEVAAVEVTVGVVTRSRSAATGKAVAPRRKRAPAGEP-----AA 57
Db 63 CRSAALDAPCTQSLQLPARSRTLQPLSGGRRPAGWDRGFPQPGVPOGHLPRAHAGAR 121
QY 58 AVSAGGGGCG-----YIHLRSRM-----LFMAPPQOPSVDSVPTPVEAADGAAG----- 103
Db 122 VHQAGGPGGRATARDRLDHGQQLSAAGAAEADGHAAPAVRQOB-GVRAAGGAAPVDPV 180

QY 104 ---OQRAALAGLSRCSSTASSVNLGLGQGRS---HTCRSY-----DAEAGGDHVL- 150
Db 181 LQLPRGAACRLGTERAGAAVFAVATQGVLAANPRHLQCRFQPEGVDPDLLAVVEHQLF 240
QY 151 -----VDVSAASNSGSGPDRRERTTSSRAHGELSDLESLLAGHKTGFSPLPAATPAEEL 205
Db 241 AGNPVAEDPPAGDERGNARLRRRPRSGERSGFLQVREHVAGR-----AVQH 289
QY 206 IVPFAHEIQEFAAAEAAQAKRFASKYN---FDFVRGVPLD 243
Db 290 VVSGTQEQROAVATVAADHDQVAVSLFRGEAVDFLARLPVD 330

RESULT 7
US-09-252-991A-30198
;; Sequence 30198, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; PRIOR FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 30198
;; LENGTH: 1239
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30198

Query Match 8.5%; Score 110.5; DB 4; Length 1239;
Best Local Similarity 23.2%; Pred. No. 0.069;
Matches 59; Conservative 25; Mismatches 93; Indels 77; Gaps 10;
QY 8 CRGAAGAEVAAVE---VTQVGVVTRSRSAATGKAVAPRRKRAPAGEPAA----- 57
Db 141 CRAAAEALDAVDKPVQALRGILQAAEQVLAQGGPGLAPLEIEREAVGLGALGGHDQR 200
QY 58 -----AVSAGDG-----GSCVIHLR-----SRMLFMAPPQOPSVDS 90
Db 201 GRIGQRHETDVGALPCVGTAGTGLAAYRSVCHHLKCGRAARAVPFSSAPAPSSARFS 260
QY 91 VPTPVEA---ADGAAGQGAALAAGL-----SRCSSSTASSVNLGLGG 129
Db 261 IAATASAMPGGAGGAGGAGKAPHGAGLVRRVAGSLDDRLGLVRSRFAELGDLVEAQAV 320
QY 130 QRSHTCRSYDAEAGGDHVLV-----DVSAASNSGSGPDRRERTTSSRA 176
Db 321 DVGSH---AVDLVRVG---HVLLHGRQFAEDRDRQLGRQCEHIGGIAQTIVREVTGRSGD 375
QY 177 HGELSDLESLLAGH 190
Db 376 HGG-TGLDASLVAH 388

RESULT 8
US-09-252-991A-27220
;; Sequence 27220, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; PRIOR FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788

```

153  QY  -----VSAASNG--SGPD--RRETTPS--SRAHGELSDLESDLA---GHKTGP 194
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
219  Db  RGRSRKVPARSFGVQGGADPRRRRGDVPFHPAQRHGRDPDLPVPALWEGDHPDRP 278
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
195  QY  SLPAATPAAEALIVP 208
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
279  Db  GLPRFRPAARLAEP 292
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
US-09-252-991A-20675
; Sequence 20675, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20675
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20675

```

```

QY      2 GKYMRCXGAAGAAEVAAVETQVVGVRTSRGAAAT-----GGVAKVAPRRKRPAAG-- 53
Db      311 GRHGPPARPAARRRAAVGEAAGAGOGPRRTGAARTQSPCSAGSAALRAHRRAEAGSAGE 370
QY      54 -----EPAAAASAGDGGSCYIHLRSMLFMAP-----POPOPSVD 89
Db      371 LRARPGEWPAAQFASAGAAPGEGRGRTARAARRQAARQRPRTAVTRDEPDQGRPEPPP--- 427
QY      90 SVPTPYEAADGAAGQGCGAALAAGLSRCSSTASSVNLGLGGQRGSHTCRSYDAAEA----- 144
Db      428 ---PFPQAQPED----QAPAAAACGDAARDP-----QRTHGTAPAEGAAQPQRQPD 472
QY      145 GGDHVILVDVSAANSNCSGDPDRER---RETPSSRAH-----GELSDLESDLACHKTGPS 195
Db      473 HGHSHQPGDPCVLPRGGARPRQPARGRHGCPGHRRLHPAARGEPAALRPRAAA--TGPA 530
QY      196 LPAATPAABE--LIVPPAIHEIQEFFAA-AEAAQAKRFASKYNFDVRGVPLDAGGRFEWA 251
Db      531 LPARTWATEEGRPVPFAEDRAYRATPADPALLRH-----PGAGGGRPAPA 577
QY      252 P 252
Db      578 P 578

```

Db 578 P 578

RESULT 11

US-09-252-991A-21386

; Sequence 21386, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18


```

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21386
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21386

```

Query Match 8.3%; Score 108.5; DB 4; Length 973;
Best Local Similarity 25.6%; Pred. No. 0.078;
Matches 75; Conservative 22; Mismatches 93; Indels 103; Gaps 13;

QY	31	SRSAATGGVAKVAPRRKRPAAGPAAVSAAGD-----GSCSYHLRSMLF	78
Db	616	ARTAAGGGYRRARLLRLLRAAGHRPAARTAAPRRGRGMPGDRHGG-----RSAGLA	669
QY	79	MAPPPQPSVDVSTPVEADGAAGQAQGAALAAGLSRCSTASSVNLGI-----GQ	130
Db	670	AAPGCPAP-----AARPGAGRPEAAGAGCELC-----DLHLGHRGAQGRGQP	714
QY	131	RGSHTCRSYDAAEAGGDHVLVDVSAASNSGG-----PDRERETTPS-----	173
Db	715	RGGDQYHRRARPAAAGERI---GSLAGLRAGLSVGLRPPRRRCQPGPAGFTGAR	771
QY	174	-----SRAHGBLSLDLESLAG-----HKTGPSLPAATPAAE	204
Db	772	CRCLGGGYPAACGEFVELGACLAGDGPAGCEPGLSOSAGGAVRLGGPGAPPPAPT	831
QY	205	LI--VPPAHEIQEFFA--AEAAQAKRFASKYNDFVR-----GVPLDAGER	247
Db	832	LCRRLLPACAGWRYRSHLVEPAERYYGAAALAFDSLPAFATAGTGLP---GGR	881

RESULT 12

```

US/09-252-991A-30345
; Sequence 30345, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND
; TITLE OF INVENTION: AERUGINOSA FOR DIA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,9
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,78
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,19
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30345
; LENGTH: 1395
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30345

```

Query Match	8.3%	Score 108.5;	DB 4;	Length 1395;
Best Local Similarity	25.4%;	Pred. No. 0.13;		
Matches 82;	Conservative 22;	Mismatches 134;	Indels 85;	Gaps 15;

Qy	6	KRCRGAAGAAVAV-----EVTQVGV---RTRSRSAATGGVAKVAPRRKRPAAGBPA	56
		:	
Db	247	RRGRG-VGPVAAAVGGDLGRDRIAIVAGGEDHRAARFRSAAAAAIAQVDRGRDRDGVHGS	305
		:	
Qy	57	AAYSAG-----GDGGS CYIHLRSMLFWAPQOPGSVDVSTPVEAA-----DGA	101
		:	
Db	306	AALAAGIAGRVSGGVDPHGSIRQRRGGEAPVAAAVGGDLDPRTAAAIICQGHGGRFAGCA	365
		:	
Qy	102	AGQGGAA-----LAAGLSRCSTAS-----SVNLGLGORGSHTCRSYDAEAAG	145
		:	
Db	366	ADDDSVARIHRRGRHRHVHGGAGRSAGIAGRVGRGVHDHGAVGQRRAR-CEAPVATAVG	424
		:	
Qy	146	GD---HVLVDVSAASNSGS-GPDRER-----RETPSSRAHGELS-----	181

[illegible]

RESULT 13

```

US-09-252-991A-17884
; Sequence 17884, Application US/09252991A
; Patent No. 6551795
;
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17884
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17884

```

Query Match 8 3% Score 108. DB 4. Length 363.

	Query: MACC1	9.38; Score 1067; DS 4; Length 362;
	Best Local Similarity	27.2%; Pred. No. 0.023;
	Matches 56; Conservative	Mismatches 93; Indels 42; Gaps 8;
QY	32	RSAATGGVAKVAPRRKRA- - - - -GEPAAVSAG-GDGSCSYIHLRSLFMAPP 82
Dd	37	QGTAASGGGCLQRRLREGPAERRI VDDGI PGAGGGAGRGGRALPDGGRGL- - -PT 92 : :
QY	83	QPOPSVDVPTVEAADGAAGCGQAALAAGLSRCSSTASSVNLLGLGCRSGHTRCSYDAA 142 : :
Dd	93	PPRPATIV- - - -AATGRASQVPARQAQAADRDGPVPRAGTGRGAGRHPPGQSRLR 147 : :
QY	143	EAGGDHVLVDVASNSGSGPDRRETTTSSRAHGELSDLSDLAGHKTPSLP-AATP 201 :
Dd	148	AA- - - - - - - - - - -GRKPAARRRPPRAGRAQAFC- - - - -LERARRAPGTFGAYP 186 :
QY	202	AALIIVPPAHEIQEFFAAAEAAQAQR 227 :
Dd	187	QA--LARPAPALQQGRAGHDRAER 210 :

RESUIT 14

```

RESOLUTION 14
US-09-252-991A-32621
; SEQUENCE 32621, APPLICATION US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32621
; LENGTH: 553

```

Job time : 34 secs

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32621

Query Match      8.3%; Score 108; DB 4; Length 553;
Best Local Similarity 26.0%; Pred. No. 0.04;
Matches 66; Conservative 22; Mismatches 82; Indels 84; Gaps 11;

QY 11 AGAEEVAEVTQVGVTRSRSAATGGVAKVAPRRKAPAGEPAAAVSAGDGGSCYI 70
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 302 MAHAALAAPDRSVGPAPADRSFAAGAH-----ARRAKKAPA-RPAAAVASGG----- 347

QY 71 HLRSMLFMAPPQPOPSVDSVPTPVEAADCAAGQQAALAGLSRCSSTASSVNLGLGQ 130
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 348 ---GLDLALA---GQPLACSGAPAAALGSGPGR---VAALRRVAPPAQPVQ-----Q 392

QY 131 RGS--HTC-----RSYDAEAGGDHVLVDVSAANSNGSGDPDRERRETTSSRAHGEL 180
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 393 RGAHRLCPRCARLRPAKDAAGGRSH-----TGTRAGAABRRRTGPGETATGDC 444

QY 181 SDLESDLAGHKTGPSILPAATPAELIVPPAHEIQEFFAAAEAAQAKRFASKYNEDFVRGV 240
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 445 QQPAAE-----ETVPAACPAHRPEQPG----- 467

QY 241 PLDAGGRFEWAPV 254
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 468 PVAAGPRRTLATV 481

RESULT 15
US-09-252-991A-22537
; Sequence 22537, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22537
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22537

Query Match      8.2%; Score 107.5; DB 4; Length 534;
Best Local Similarity 27.2%; Pred. No. 0.043;
Matches 58; Conservative 18; Mismatches 78; Indels 59; Gaps 12;

QY 9 RGAAGAEEVAEVTQVGVTRSRSAATGGVAKVAPRRKRA-----PAGEPAAAVSAGG 63
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 127 RGTA-VQPRILVALRAVAEPTAGGAHRTG-----RRRRADDAGGTAGDP-ARLPAPG 177

QY 64 DGGSCVTHLRSLFMVAPQPSVDSVPTPVEAADCAAGQQAALAGLSRCSSTASSV 123
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 178 PGAGAQLRHHSR-----PARP-----AGTADAGLGGVRLIA----- 210

QY 124 NLGLGGQSGSHTCR--SYDAEAGGDHVLVDVSAANSNGSGDPDRERRETTSS--SRAHGEL 180
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 211 -LDLPQAGRPARLEGGDEADAG-----PAQGAEPFRQHR--LPALRLRHG-- 255

QY 181 SDLESDLAGHKTGPSILPAATPAELIVPPAHEI 213
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 256 ADLHRPGRTERAAPVAPFARGPAADRRTGPAHRL 288
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:53:56 ; Search time 130 Seconds
(without alignments)
633.697 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 1304

Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVLDAAGRFEWAPVSI 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1304	100.0	256	9	US-09-993-308-2
2	1304	100.0	256	11	US-09-993-808B-2
3	1217.5	93.4	263	12	US-10-425-114-63204
4	1059.5	81.2	221	12	US-10-425-114-69245
5	669	51.3	262	15	US-10-333-006-10
6	669	51.3	417	16	US-10-437-963-195115
7	668.5	51.3	278	12	US-10-425-114-71211
8	665	51.0	227	9	US-09-993-308-6
9	665	51.0	227	11	US-09-993-808B-6
10	399.5	30.6	116	15	US-10-333-006-14
11	315.5	24.2	354	16	US-10-437-963-122432
12	314	24.1	423	16	US-10-437-963-122443
13	279	21.4	226	15	US-10-333-006-55
14	238	18.3	194	16	US-10-437-963-120870
15	235	18.0	194	15	US-10-333-006-44

16	231	17.7	255	12	US-10-425-114-61054	Sequence 61054, A
17	226	17.3	218	16	US-10-437-963-128205	Sequence 128205, A
18	220.5	16.9	225	16	US-10-437-963-198574	Sequence 198574, A
19	218	16.7	205	12	US-10-424-599-235800	Sequence 235800, A
20	215.5	16.5	222	16	US-10-451-139-21	Sequence 21, Appl
21	215	16.5	248	12	US-10-425-114-59718	Sequence 59718, A
22	212.5	16.3	222	12	US-10-688-291-4	Sequence 4, Appli
23	205.5	15.8	224	12	US-10-424-599-210190	Sequence 210190, A
24	200	15.3	190	9	US-09-993-308-4	Sequence 4, Appli
25	200	15.3	190	11	US-09-993-808B-4	Sequence 4, Appli
26	178	13.7	196	9	US-09-733-507-16	Sequence 16, Appl
27	178	13.7	196	16	US-10-451-139-15	Sequence 15, Appl
28	178	13.7	289	16	US-10-451-139-17	Sequence 17, Appl
29	160	12.3	205	12	US-10-424-599-182928	Sequence 182928, A
30	153	11.7	85	16	US-10-767-703-37255	Sequence 37255, A
31	150.5	11.5	176	12	US-10-424-599-212181	Sequence 212181, A
32	149	11.4	327	12	US-10-688-291-34	Sequence 34, Appl
33	145	11.1	137	16	US-10-451-139-13	Sequence 13, Appl
34	144.5	11.1	108	15	US-10-333-006-16	Sequence 16, Appl
35	142.5	10.9	137	9	US-09-733-507-14	Sequence 14, Appl
36	134	10.3	90	15	US-10-333-006-13	Sequence 13, Appl
37	134	10.3	196	16	US-10-451-139-22	Sequence 22, Appl
38	134	10.3	212	9	US-09-733-507-12	Sequence 12, Appl
39	134	10.3	212	16	US-10-451-139-11	Sequence 11, Appl
40	132.5	10.2	180	12	US-10-425-114-49866	Sequence 49866, A
41	132	10.1	95	16	US-10-767-701-36263	Sequence 36263, A
42	129.5	9.9	93	15	US-10-333-006-15	Sequence 15, Appl
43	129.5	9.9	204	16	US-10-437-963-169561	Sequence 169561, A
44	127.5	9.8	175	12	US-10-425-114-46534	Sequence 46534, A
45	127.5	9.8	176	12	US-10-425-114-61692	Sequence 61692, A

ALIGNMENTS

RESULT 1

US-09-993-308-2
; Sequence 2, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR FILING DATE: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Zea mays
US-09-993-308-2

Query Match 100.0%; Score 1304; DB 9; Length 256;
Best Local Similarity 100.0%; Pred. No. 1e-101;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKYMRCRGAAGAEVAEVTQVGVRTSRSAATGGVAKVAPRRKRAPAGEPAAVS 60
DB 1 MGKYMRCRGAAGAEVAEVTQVGVRTSRSAATGGVAKVAPRRKRAPAGEPAAVS 60

QY 61 AGDGGSCYIHLSRMLFWAPQPSVDSVPTFEAAGAACQOAGALAAAGSRCSSTA 120
DB 61 AGDGGSCYIHLSRMLFWAPQPSVDSVPTFEAAGAACQOAGALAAAGSRCSSTA 120

QY 121 SSVNLGLGQGRSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRRTTPTSSRAHGE 180

Db 121 SSVNLGLGQSGHTCRSYDAEAGGHHVLDVSAASNSGSGPDRERETTPSSRAHGE 180
Qy 181 SDLESDLAGHTGSLPAATPAELIIVPPAHEIQEFPFAAAEAAQAKRFASKYNFDFVRGV 240
Db 181 SDLESDLAGHTGSLPAATPAELIIVPPAHEIQEFPFAAAEAAQAKRFASKYNFDFVRGV 240
Qy 241 PLDAGGRFEWAPVSI 256
Db 241 PLDAGGRFEWAPVSI 256

RESULT 2

US-09-993-808B-2
; Sequence 2, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 256
; TYPE: PRT
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(256)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-993-808B-2

Query Match 100.0%; Score 1304; DB 11; Length 256;
Best Local Similarity 100.0%; Pred. No. 1e-101;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGKYMRCRGAAGAEEVAEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGKYMRCRGAAGAEEVAEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
Qy 61 AGDGGGSCYIHLRSMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSCSSTA 120
Db 61 AGDGGGSCYIHLRSMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSCSSTA 120
Qy 121 SSVNLGLGQSGHTCRSYDAEAGGHHVLDVSAASNSGSGPDRERETTPSSRAHGE 180
Db 121 SSVNLGLGQSGHTCRSYDAEAGGHHVLDVSAASNSGSGPDRERETTPSSRAHGE 180
Qy 181 SDLESDLAGHTGSLPAATPAELIIVPPAHEIQEFPFAAAEAAQAKRFASKYNFDFVRGV 240
Db 181 SDLESDLAGHTGSLPAATPAELIIVPPAHEIQEFPFAAAEAAQAKRFASKYNFDFVRGV 240
Qy 241 PLDAGGRFEWAPVSI 256
Db 241 PLDAGGRFEWAPVSI 256

RESULT 3

US-10-425-114-63204
; Sequence 63204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63204
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMOL7355F05_FLI1.pep
US-10-425-114-63204

Query Match 93.4%; Score 1217.5; DB 12; Length 263;
Best Local Similarity 95.3%; Pred. No. 2.1e-94;
Matches 245; Conservative 3; Mismatches 4; Indels 5; Gaps 3;
Qy 1 MGKYMRCRGAAGAEEVAEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAV 59
Db 11 MGKYMRCRGAAGAEEVAEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAV 70
Qy 60 SAGDGGGSCYIHLRSMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSCSST 119
Db 71 CAGDGGGSCYIHLRSMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSCSST 128
Qy 120 ASSVNLGLGQSGHTCRSYDAEAGGHHVLDVSAASNSGSGPDRERETTPSSRAHGE 179
Db 129 ASSVH--VGGQSGHTCRSDDAEAGGHHVLDVSAASNSGSGPDRERETTPSSRAHGE 186
Qy 180 LSDLESDLAGHTGSLPAATPAELIIVPPAHEIQEFPFAAAEAAQAKRFASKYNFDFVRG 239
Db 187 LSDLESDLAGHTGSLPAATPAELIIVPPAHEIQEFPFAAAEAAQAKRFASKYNFDFVRG 246
Qy 240 VPLDAGGRFEWAPVSI 256
Db 247 VPLDAGGRFEWAPVSI 263

RESULT 4

US-10-425-114-69245
; Sequence 69245, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69245
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI1.pep
US-10-425-114-69245

Query Match 81.2%; Score 1059.5; DB 12; Length 221;
Best Local Similarity 94.7%; Pred. No. 3.4e-81;
Matches 213; Conservative 3; Mismatches 4; Indels 5; Gaps 3;
Qy 33 SAAATGGVAK-VAPRRKRAPAGEPAAAVSAGDGGGSCYIHLRSMLFMAPPQPSVDSV 91
Db 1 SAAATGGVAKVAPRRKRAPAGEPAAAVSAGDGGGSCYIHLRSMLFMAPPQPSVDSV 58

QY 92 PTPVEADCAAGQQAALAAAGLSCSTASSVNLGLGQGRSHTCRSYDAAEAGGDHVLV 151
Db 59 PTPAEADCAAGQQAALAAAGLSCSTASSVH--VGGQSGHTCRSDAAEAGGDHVLV 116
QY 152 DVSAASNSGSPDRERRETTSSRAHGLSLESDLAGHKTGPSLPAATPAELIIVPPAH 211
Db 117 DVSAASNSGSPDRERRETTSSRAHGLSLESDLAGHKTGPSLPAATPAELIIVPPAH 176
QY 212 EI0FFFAAAEAQAARFASKYNFDFVRGVPLDAGGRFEPWVSI 256
Db 177 EI0FFFAAAEAQAARFASKYNFDFVRGVPLDAGGRFEPWVSI 221

RESULT 5

US-10-333-006-10
; Sequence 10, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A11le S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-333-006-10

Query Match 51.3%; Score 669; DB 15; Length 262;
Best Local Similarity 60.3%; Pred. No. 3.2e-48;
Matches 161; Conservative 19; Mismatches 67; Indels 20; Gaps 10;

QY 1 MGKYMRCRCGAGAEVAAEVTVGVVTRSRSAATGG-----VAKVAPRRKRAPAGE 54
Db 1 MGKYMRCRCGAGAEVTVGVVTRSRSAATGG-----VAKVAPRRKRAPAGE 60
QY 55 PAAAVSAGDGGSCYTHLRSMFLMAPPQPOPQSDVSVTPVEADCAAGQQAALAAAGLS 114
Db 61 TAVVGTTRDGGSCYQLRSMFLMAPPQPOPQSDVSVTPVEADCAAGQQAALAAAGLS 118
QY 115 RCSSTASSVNLGLGQGRGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERRETTSS 174
Db 119 RCSSTASSVD--AAAQDRSLACRS-DVAEAGSEH--VPEGSASDSASGRDRERRETTSS 173
QY 175 RAHGELSDLESILA-GHKTGPSLP--RATPAELI-----VPPAHEIOEFFFFAAAEAAQKRF 228
Db 174 FLPGVEVSDLESILAGGQKRSRPLPSAATAQAOTRKPPIPPAAEIEIAFFAAAEAEAKRF 233
QY 229 ASKYNDFVRGVPLDAGGRFEPWVSI 255
Db 234 AAKYNDFVRGVPLDA-GRFEWTPVVS 259

RESULT 6

US-10-437-963-195115
; Sequence 195115, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195115
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1.pep
US-10-437-963-195115

Query Match 51.3%; Score 669; DB 16; Length 417;
Best Local Similarity 60.3%; Pred. No. 5.6e-48;
Matches 161; Conservative 19; Mismatches 67; Indels 20; Gaps 10;
QY 1 MGKYMRCRCGAGAEVAAEVTVGVVTRSRSAATGG-----VAKVAPRRKRAPAGE 54
Db 156 MGKYMRCRCGAGAEVTVGVVTRSRSAATGG-----VAKVAPRRKRAPAGE 215
QY 55 PAAAVSAGDGGSCYTHLRSMFLMAPPQPOPQSDVSVTPVEADCAAGQQAALAAAGLS 114
Db 216 TAVVGTTRDGGSCYQLRSMFLMAPPQPOPQSDVSVTPVEADCAAGQQAALAAAGLS 273
QY 115 RCSSTASSVNLGLGQGRGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERRETTSS 174
Db 274 RCSSTASSVD--AAAQDRSLACRS-DVAEAGSEH--VPEGSASDSASGRDRERRETTSS 328
QY 175 RAHGELSDLESILA-GHKTGPSLP--RATPAELI-----VPPAHEIOEFFFFAAAEAAQKRF 228
Db 329 FLPGVEVSDLESILAGGQKRSRPLPSAATAQAOTRKPPIPPAAEIEIAFFAAAEAEAKRF 388
QY 229 ASKYNDFVRGVPLDAGGRFEPWVSI 255
Db 389 AAKYNDFVRGVPLDA-GRFEWTPVVS 414

RESULT 7

US-10-425-114-71211
; Sequence 71211, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71211
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3078-022-G5_FLI.pep
US-10-425-114-71211

Query Match 51.3%; Score 668.5; DB 12; Length 278;
Best Local Similarity 73.1%; Pred. No. 3.8e-48;

Matches 174; Conservative 7; Mismatches 30; Indels 27; Gaps 17;
QY 1 MGKMKRCRGAAGAEVAAVEVTQVGVTRSRSAAT-GGVAKVA-PRRKA--PA---- 52
Db 47 MGKMKRRRGAAGAEVAAVEVQVGVTRSRSAATGGGVAKVAPRRKKALLPAANET 106
QY 53 --GEPAAVSAAGDGGG-CYHLRSMLFMAPPQPSVDSVPTVEAADGAAGQOQ-AA 108
Db 107 ASGEPGAVGAGGCGGSCCYHLRSMLFMAPPQOQS--AALTPEAA-GAAXQGGVVA 163
QY 109 LAAGLSCSSTASSVNLGLGQOR---GSHTCRSYDA-AEAGGDHVLVDVSAASNSGSGPD 164
Db 164 LAAGLSCSSTASSVD--VGQQPASGSHACKSDAAPAEVGDHV-PDVVTASNSGSPD 220
QY 165 RERETTP-SSRAH-GELSDLESdiag-HKTG-PSLPA-ATPAELIYPPAHEIOEFF 217
Db 221 RERETTPSSRAHGGELSDLESdiagVGRQKTCSSSPATTTSAELIYPPAHEIOEFF 278

RESULT 8

US-09-993-308-6
; Sequence 6, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: The 'Xaa' at location 111 stands for Gln.
; US-09-993-308-6

Query Match 51.0%; Score 665; DB 9; Length 227;
Best Local Similarity 72.6%; Pred. No. 5.8e-48;
Matches 172; Conservative 6; Mismatches 31; Indels 28; Gaps 16;
QY 1 MGKMKRCRGAAGAEVAAVEVTQVGVTRSRSAAT-GGVAKVA-PRRKA--PA---- 52
Db 1 MGKMKRRRGAAGAEVAAVEVQVGVTRSRSAATGGGVAKVAPRRKKALLPAANVT 60
QY 53 --GEPAAVSAAGDGGG-CYHLRSMLFMAPPQPSVDSVPTVEAADGAAGQOQ-AA 108
Db 61 TSGEPAVAGAGGCGGSCCYHLRSMLFMAPPQOQS--AALTPEAA-GAAXQGGVVA 117
QY 109 LAAGLSCSSTASSVNLGLGQORSHCTCRSYDA-AEAGGDHVLVDVSAASNSGSGPD 167
Db 118 LAAGLSCSSTASSVDV-----GGHACRSDAAPAEVGDHV-PDVVTASNSGSPDR 170
QY 168 RERETTP-SSRAH-GELSDLESdiag-HKTG-PSLPA-ATPAELIYPPAHEIOEFFAA 219
Db 171 RERETTPSSRAHGGELSDLESdiagVGRQKTCSSSPATTTSAELIYPPAHEIOEFFAA 227

RESULT 9

US-09-993-808B-6
; Sequence 6, Application US/09993808B
; Patent No. US20040003433A1
; GENERAL INFORMATION:

; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(227)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-993-808B-6

Query Match 51.0%; Score 665; DB 11; Length 227;
Best Local Similarity 72.6%; Pred. No. 5.8e-48;
Matches 172; Conservative 6; Mismatches 31; Indels 28; Gaps 16;
QY 1 MGKMKRCRGAAGAEVAAVEVTQVGVTRSRSAAT-GGVAKVA-PRRKA--PA---- 52
Db 1 MGKMKRRRGAAGAEVAAVEVQVGVTRSRSAATGGGVAKVAPRRKKALLPAANVT 60
QY 53 --GEPAAVSAAGDGGG-CYHLRSMLFMAPPQPSVDSVPTVEAADGAAGQOQ-AA 108
Db 61 TSGEPAVAGAGGCGGSCCYHLRSMLFMAPPQOQS--AALTPEAA-GAAXQGGVVA 117
QY 109 LAAGLSCSSTASSVNLGLGQORSHCTCRSYDA-AEAGGDHVLVDVSAASNSGSGPD 167
Db 118 LAAGLSCSSTASSVDV-----GGHACRSDAAPAEVGDHV-PDVVTASNSGSPDR 170
QY 168 RERETTP-SSRAH-GELSDLESdiag-HKTG-PSLPA-ATPAELIYPPAHEIOEFFAA 219
Db 171 RERETTPSSRAHGGELSDLESdiagVGRQKTCSSSPATTTSAELIYPPAHEIOEFFAA 227

RESULT 10

US-10-333-006-14
; Sequence 14, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatfield, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)

OTHER INFORMATION: Xaa = any amino acid
US-10-333-006-14

Query Match 30.6%; Score 399.5; DB 15; Length 116;
Best Local Similarity 81.9%; Pred. No. 5.8e-26;
Matches 95; Conservative 0; Mismatches 14; Indels 7; Gaps 7;

QY 146 GDHVLVDVSAASNSGSGPDRERETTP--SSRAH--GELSDLESIDAG--HKTG--PSLPA--AT 200
Db 3 GDHV--PDVXASNSGSGVPDRERETTPSSRAHGGELSDLESIDLVGKQTCSSSPATT 61
QY 201 PAELIVPPAHBIOEFPFAAAEAQAKFASKYNFDFVRGVPLDAGGPFEPWAPVVISI 256
Db 62 SAAELIVPPAQEIQEFPFAAAEAHAKFASKYNFDFVRGVPLDAG--GRFETWPGVSI 116

RESULT 11

US-10-437-963-122432

Sequence 122432, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 122432

LENGTH: 354

TYPE: PRP

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_2535C.1.p

US-10-437-963-122432

Query Match 24.2%; Score 315.5; DB 16; Length 354;
Best Local Similarity 37.4%; Pred. No. 2.7e-18;
Matches 105; Conservative 26; Mismatches 81; Indels 69; Gaps 13;

QY 1 MGKMKRCGAAGAEEVAEEVTVGVTRRSAAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGK-KKRDGAARRQARVVVG---GVRTR-----AAVTARRVVASAEEGCGLVG 46
QY 61 AGDGGG-----CVIHLRSMLFMAPPPQPPQSVDSVPTPVEAADGAAGQQAALAA 112
Db 47 RGGGGSGDGDGEGGCGYLRLRRRL-----PFVAAVSSRREELGDS 90
QY 113 LSRCSSTASSVNLGLGQSGHTCRSYD--AAEAGGDHVLVDVSAASNSGSGPDRERETT 171
Db 91 VAEAASSSSSRRAVELLGCSEEEAEAKVCTQAGEDH--DEESSVDSGCG--RERSATT 146
QY 172 PSSR---ANGELSDLESID-----LAGHKTGSPSLPATPAELIVPP--AH 211
Db 147 PSSRRPPGDADSDAESNQAQCMCRSSSTTSAAPAFHAG---ATTFRFMWAPFAAAA 202
QY 212 EQEFPFAAAEAQAKFASKYNFDFVRGVPLDAG--GREW 250
Db 203 EIEEFLAAAESEAEERFAAKYNFDFVRGVPLDAGGAGREFW 243

RESULT 12

US-10-963-122443

Sequence 122443, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 122443
LENGTH: 423
TYPE: PRP
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_2536C.1.p
US-10-437-963-122443

Query Match 24.1%; Score 314; DB 16; Length 423;
Best Local Similarity 36.4%; Pred. No. 4.5e-18;

Matches 107; Conservative 27; Mismatches 78; Indels 82; Gaps 14;

QY 1 MGKMKRCGAAGAEEVAEEVTVGVTRRSAAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGK-KKRDGAARRQARVVVG---GVRTR-----AAVTARRVVASAEEGCGLVG 46
QY 61 AGDGGG-----CVIHLRSMLFMAPPPQPPQSVDSVPTPVEAADGAAGQQAALAA 112
Db 47 RGGGGSGDGDGEGGCGYLRLRRRL-----PFVAAVSSRREELGDS 90
QY 113 LSRCSSTASSVNLGLGQSGHTCRSYDAA-----EAGGDHVLVDVSAASN 158
Db 91 VAEAASSSSSRRAVELLGCSEEEAEAKVIDEAAWVVLQVCTQAGEDH--DEESSVGD 148
QY 159 SGSGPDRERETTPSSR---ANGELSDLESID-----LAGHKTGSPSLPAT 200
Db 149 SGCG--RERSATTTPSSRRPPGDADSDAESNQAQCMCRSSSTTSAAPAFHAG---ATT 202
QY 201 PAELIVPP--AHEIOEFPFAAAEAQAKFASKYNFDFVRGVPLDAG--GREW 250
Db 203 RSFRFMWAPFAAAAEIEEFLAAAESEAEERFAAKYNFDFVRGVPLDAGGAGREFW 256

RESULT 13

US-10-333-006-55

Sequence 55, Application US/10333006

Publication No. US20040019926A1

GENERAL INFORMATION:

APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.

APPLICANT: Peres Bota, Adrian Marius

APPLICANT: Droual, Anne-Marie

APPLICANT: Mironov, Vladimir

APPLICANT: Inz, Dirk

APPLICANT: Hatzfeld, Yves

TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
FILE REFERENCE: 1187-13

CURRENT APPLICATION NUMBER: US/10/333,006

CURRENT FILING DATE: 2003-01-14

PRIOR APPLICATION NUMBER: PCI/IB01/01492

PRIOR FILING DATE: 2001-06-29

PRIOR APPLICATION NUMBER: US 60/218,471

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/241,219

PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn version 3.1

SEQ ID NO 55

LENGTH: 226

TYPE: PRP

```
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: probe or primer
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119)..(119)
; OTHER INFORMATION: Xaa = any amino acid
US-10-333-006-55

Query Match                21.4%; Score 279; DB 15; Length 226;
Best Local Similarity      34.4%; Pred. No. 1.8e-15;
Matches 98; Conservative 26; Mismatches 71; Indels 90; Gaps 13;

QY 1 MGKYMRCRGAAGAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
DB 1 MGK-KKRDGGAARQARVVG-----AAVTARRVVAEEGCGLVG 46
QY 61 AGDGGG-----CYHLRSRLFMAPPQPSVDSVPTPVEAADGAAGQQAALAAAG 112
DB 47 RGGGGGGDDEGGCYLRLRSRL-----PFVAAAVVSSRREALGDS 90
QY 113 LSRCSSTASSVNLGLGGQSGHTCRSYDAEAGDGHVLDVSAASNSGSGPDRRRTTP 172
DB 91 VAESAASSSSRAVELLG-----C-----SGEEAAMAEKXSA-----TTP 124
QY 173 SSR---AHGELSDLES-----LAGHKTGSLPAATPAALIVPP---AHE 212
DB 125 SSRPPGDADSDAESQEAQKQCMCRSSSTTSAAAFHAG-----ATTSFRMAPPAAAE 180
QY 213 IQEFAAAEAQAARFASKYNFDFVRGVPLDAG--GRFEMAPVVS 255
DB 181 IEELFAAERSEAFKYNFDFVRGVPLDAGGAGREFWAVGS 225

RESULT 14
US-10-437-963-120870
; Sequence 120870, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: Ia Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 120870
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MKT4530_23950C.1.pep
US-10-437-963-120870

Query Match                18.3%; Score 238; DB 16; Length 194;
Best Local Similarity      31.5%; Pred. No. 4.3e-12;
Matches 82; Conservative 22; Mismatches 76; Indels 80; Gaps 11;

QY 1 MGKYMRCRGAAGAAVEVTQV-----VGVTRSRSAATGGVAKVAPRRKRAPAGEPA 56
DB 1 MGKYMRCRGAAGAAVEVTQV-----VGVTRSRSAATGGVAKVAPRRKRAPAGEPA 56
QY 57 AAVSAGGCGGSCYIHLRSRLFMAPPQPSVDSVPTPVEAADGAAGQQAALAAAGLSC 116
DB 44 ----QGGE-----YLELRSRLEKLP-----PPRRRATAAA-----ATA 78
QY 117 SSTATSVNLGLGGQSGHTCRSYDAEA-----GDHVLVDVSAASNSGSGPDRRRTTP 172
DB 79 DATAT-----ESAEAEVSGGENVL-ELEAM-----ERNTRRTTP 112
QY 173 SS--RAHGELSDLESDLAGHKTGSLPAATPAALIVPPAHEIOEFFFAAAQAARFAS 230
DB 113 CSLIRDPDTISTPGSTTRRSHSSCHKVQTPVRHNIIPASAELEAFPAEEQRORQAFID 172
QY 231 KYNFDFVRGVPLDAGGRFEW 250
DB 173 KYNFDFVNDCELP--GRFEW 190

Search completed: October 1, 2004, 17:05:50
Job time : 131 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 17:03:42 ; Search time 41 Seconds
(without alignments)
600.611 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256

Sequence: 1 MCKYMKCRGAAGAEVAAVE.....VRGVPLDAGRFEWAPVSI 256

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 78: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	3.9	277	H84314	cytochrome aa3 con
2	8	3.1	173	T45993	hypothetical prote
3	8	3.1	181	T45990	hypothetical prote
4	8	3.1	215	T03680	plasma membrane pr
5	8	3.1	272	S41498	DNA excision repai
6	8	3.1	334	T29061	probable integral
7	8	3.1	404	D70977	hypothetical prote
8	8	3.1	427	AC1627	N-acetylmuramoyl-L
9	8	3.1	427	AI1264	N-acetylmuramoyl-L
10	8	3.1	452	S03829	nifB protein - Rho
11	8	3.1	599	S65180	hypothetical prote
12	8	3.1	601	S12004	tyramine receptor
13	8	3.1	601	JH0170	octopamine recepto
14	8	3.1	1306	A70934	hypothetical glyci
15	7	2.7	44	EXSA1	antibacterial prot
16	7	2.7	73	E70833	hypothetical prote
17	7	2.7	97	T18130	hypothetical prote
18	7	2.7	99	S59145	NADH2 dehydrogenas
19	7	2.7	101	T26641	hypothetical prote
20	7	2.7	118	S47024	ribosomal protein
21	7	2.7	118	AF1260	hypothetical prote
22	7	2.7	118	AH1622	hypothetical prote
23	7	2.7	125	H70659	hypothetical prote
24	7	2.7	125	F72785	hypothetical prote
25	7	2.7	126	F84299	hypothetical prote
26	7	2.7	130	B84300	hypothetical prote
27	7	2.7	131	B84598	hypothetical prote
28	7	2.7	132	T35141	hypothetical prote
29	7	2.7	134	PWQFE	H+-transporting tw

30	7	2.7	139	2	A26892	Mopa box protein -
31	7	2.7	140	2	S27658	hypothetical prote
32	7	2.7	142	2	T46542	hypothetical prote
33	7	2.7	142	2	S50662	hypothetical prote
34	7	2.7	152	1	MYPN	myoglobin - empero
35	7	2.7	153	2	JT0636	myoglobin - whiske
36	7	2.7	153	2	JC5190	myoglobin - rhinoc
37	7	2.7	153	2	JC7789	myoglobin - Japane
38	7	2.7	153	2	JC7791	myoglobin - thick-
39	7	2.7	160	2	B70718	hypothetical prote
40	7	2.7	163	2	E75416	hypothetical prote
41	7	2.7	163	2	F84651	probable kinetecho
42	7	2.7	163	2	B86335	hypothetical prote
43	7	2.7	185	2	B70072	conserved hypotet
44	7	2.7	187	2	H81347	probable decarboxy
45	7	2.7	192	2	T26386	hypothetical prote

ALIGNMENTS

RESULT 1

H84314 cytochrome aa3 controlling protein [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: H84314

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabori Gung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Bennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lie

A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: H84314

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-277 <STO>

A;Cross-references: GB:AE004437; NID:gi0581096; PIDN:AAG19884.1; GSPDB:GN00138

C;Genetics:

A;Gene: ccp

Query Match 3.9%; Score 10; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 0.25; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0;

QY 32 R5AAATGGVA 41

Db 216 R5AAATGGVA 225

RESULT 2

T45993 hypothetical protein F9D24.240 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C;Accession: T45993

R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23011

A;Accession: T45993

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-173 <DAN>

A;Cross-references: EMBL:AL137081

A;Experimental source: cultivar Columbia; BAC clone F9D24

C;Genetics:

A;Map position: 3

A;Introns: 41/3

A;Note: F9D24.240

Query Match 3.1%; Score 8; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
 |||||
 Db 148 LSDLES DL 155

RESULT 3
 T45990
 hypothetical protein F9D24.210 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 28-Jul-2000
 C:Accession: T45990
 R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23011
 A:Accession: T45990
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-181 <DNA>
 A:Cross-references: EMBL:AL137081
 A:Experimental source: cultivar Columbia; BAC clone F9D24
 C:Genetics:
 A:Map position: 3
 A:Introns: 47/3
 A:Note: F9D24.210
 C:Superfamily: Arabidopsis thaliana hypothetical protein F9D24.210

Query Match 3.1%; Score 8; DB 2; Length 181;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
 |||||
 Db 154 LSDLES DL 161

RESULT 4
 T03680
 plasma membrane protein - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
 C:Accession: T03680
 R:Logan, D.C.; Domergue, O.; Teyssendier de la Serve, B.; Rossignol, M.
 submitted to the EMBL Data Library, October 1996
 A:Description: A new family of plasma membrane intrinsic polypeptides differentially reg
 A:Reference number: Z15001
 A:Accession: T03680
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-215 <LOG>
 A:Cross-references: EMBL:Y08609
 A:Experimental source: strain xanthi N.C; tissue-type leaf

Query Match 3.1%; Score 8; DB 2; Length 215;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 PAATPAE 204
 |||||
 Db 176 PAATPAE 183

RESULT 5
 S41498
 DNA excision repair protein XPAC - mouse
 N:Alternate names: Xeroderma pigmentosum group A-complementing protein
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S41498
 R:van Oostrom, C.T.M.; de Vries, A.; Verbeek, S.J.; van Kreijl, C.F.; van Steeg, H.
 Nucleic Acids Res. 22, 11-14, 1994
 A:Title: Cloning and characterization of the mouse XPAC gene.

A:Reference number: S41498; MUID:94173654; PMID:8127648
 A:Accession: S41498
 A:Status: preliminary; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-272 <VAN>
 A:Cross-references: EMBL:X74351; NID:9440564; PIDN:CAA52393.1; PID:9440565
 C:Superfamily: DNA excision repair protein XPAC
 C:Keywords: DNA binding; zinc finger

Query Match 3.1%; Score 8; DB 1; Length 272;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGVA 41
 |||||
 Db 53 AAATGGVA 50

RESULT 6
 T29061
 probable integral membrane protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C:Accession: T29061
 R:Redenbach, M.; Kieser, H.M.; Denapaita, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
 Mol. Microbiol. 21, 77-96, 1996
 A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb st
 A:Reference number: Z20556; MUID:97000351; PMID:8843436
 A:Accession: T29061
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-334 <RED>
 A:Cross-references: EMBL:AL031124; NID:el312893; PID:el312905; PIDN:CAA19979.1
 C:Genetics:
 A:Note: SC1C2.12c

Query Match 3.1%; Score 8; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 GAAGAEVA 17
 |||||
 Db 141 GAAGAEVA 148

RESULT 7
 D70977
 hypothetical protein RV3446C - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: D70977
 R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whithead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: D70977
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-404 <COL>
 A:Cross-references: GB:Z95389; GB:AL123456; NID:g3242261; PIDN:CAB08676.1; PID:e315967;
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3446C

Query Match 3.1%; Score 8; DB 2; Length 404;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 AGAEVA 19
 |||||

Db 131 AGAEVAAV 138

RESULT 8

AC1627
N-acetylmuramoyl-L-alanine amidase homolog lin1556 [imported] - *Listeria innocua* (strain C;Species: *Listeria innocua*
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AC1627
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative Genomics of *Listeria* species
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96787.1; PID:gl6414043; GSPDB:GN00178
A;Experimental source: strain Clapl1262
C;Genetics:
A;Gene: lin1556

Query Match 3.1%; Score 8; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 DVSAASNS 159

|||||

Db 92 DVSAASNS 99

RESULT 9

AI1264
N-acetylmuramoyl-L-alanine amidase homolog lmo1521 [imported] - *Listeria monocytogenes* C;Species: *Listeria monocytogenes*
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI1264
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative Genomics of *Listeria* species
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1264
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99599.1; PID:gl6410950; GSPDB:GN00177
A;Experimental source: strain EGB-e
C;Genetics:
A;Gene: lmo1521

Query Match 3.1%; Score 8; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 DVSAASNS 159

|||||

Db 92 DVSAASNS 99

RESULT 10

S03829
nifB protein - *Rhodobacter capsulatus* C;Species: *Rhodobacter capsulatus*
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jul-2000
C;Accession: S03829; A36914

R;Maspehl, B.; Klipp, W.; Puehler, A.

Mol. Gen. Genet. 212, 27-37, 1988

A;Title: Genetic characterization and sequence analysis of the duplicated nifA/nifB gene

A;Reference number: S03828; MUID:88232432; PMID:2836706

A;Accession: S03829

A;Molecule type: DNA

A;Residues: 1-452 <MAS>

A;Cross-references: EMBL:X07567; NID:gs748652; PIDN:CAA30450.1; PID:gs46063

R;Wang, G.; Angermuller, S.; Klipp, W.

J. Bacteriol. 175, 3031-3042, 1993

A;Title: Characterization of *Rhodobacter capsulatus* genes encoding a molybdenum transport

A;Reference number: A36914; MUID:93259949; PMID:8491722

A;Accession: A36914

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 381-452 <WAN>

A;Note: sequence extracted from NCBI backbone (NCBIN:131915, NCBIP:131916)

C;Genetics:

A;Gene: nifB

C;Superfamily: Rhizobium nifB protein

Query Match 3.1%; Score 8; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AAAAQAQA 225

|||||

Db 316 AAAAQAQA 323

RESULT 11

S65180

hypothetical protein YPL169c - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein P2520

C;Species: *Saccharomyces cerevisiae*

C;Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 19-Apr-2002

C;Accession: S65180; S69430

R;Purnelle, B.; Coster, F.; Goffeau, A.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S65154

A;Accession: S65180

A;Molecule type: DNA

A;Residues: 1-599 <PUR>

A;Cross-references: EMBL:Z73525; NID:gl370356; PID:gl370357; MIPS:YPL169c

A;Experimental source: strain S288C (AB972)

R;Purnelle, B.; Comblez, S.; Coster, F.; Naveau, F.; Goffeau, A.

submitted to the EMBL Data Library, March 1996

A;Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies a
logue to the human phosphotyrosyl phosphatase activator PTPA and a homologue to the plant

A;Reference number: S69428

A;Accession: S69430

A;Molecule type: DNA

A;Residues: 1-599 <PUW>

A;Cross-references: EMBL:X96770; NID:gl403537; PID:gl403540

C;Genetics:

A;Gene: SGD:MEX67

A;Cross-references: SGD:S0006090

A;Map position: 16L

Query Match 3.1%; Score 8; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 PPQPQPSV 88

|||||

Db 506 PPQPQPSV 513

RESULT 12

S12004

tyramine receptor - fruit fly (*Drosophila* sp.)C;Species: *Drosophila* sp.

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 04-Sep-1998

C;Accession: S12004
R;Saudou, F.; Amlaiky, N.; Plassat, J.L.; Borrelli, E.; Hen, R.
EMBO J. 9, 3611-3617, 1990
A;Title: Cloning and characterization of a Drosophila tyramine receptor.
A;Reference number: S12004; MUID:91006061; PMID:2170118
A;Accession: S12004
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-601 <SAU>
C;Superfamily: octopamine receptor type I

Query Match 3.1%; Score 8; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 33; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 61 AGDGGGSC 68
|||||
Db 435 AGDGGGSC 442

RESULT 13
JH0170
octopamine receptor type I - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 10-Dec-1999
C;Accession: JH0170
R;Arakawa, S.; Gocayne, J.D.; McCombie, W.R.; Urquhart, D.A.; Hall, L.M.; Fraser, C.M.;
Neuron 4, 343-354, 1990
A;Title: Cloning, localization, and permanent expression of a Drosophila octopamine rece
A;Reference number: JH0170; MUID:90198291; PMID:2156539
A;Accession: JH0170
A;Molecule type: mRNA
A;Residues: 1-601 <ARA>
A;Cross-references: GB:M60789; GB:M26181; NID:G158002; PIDN:AAA28731.1; PID:G158003
A;Experimental source: strain oregon-R
A;Note: this sequence shows homology with alpha-2 adrenergic receptors
C;Genetics:
A;Gene: FlyBase:Ocr
A;Cross-references: FlyBase:FBgn0004514
C;Superfamily: octopamine receptor type I
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;110-133/Domain: transmembrane #status predicted <TM1>
F;147-169/Domain: transmembrane #status predicted <TM2>
F;184-207/Domain: transmembrane #status predicted <TM3>
F;227-250/Domain: transmembrane #status predicted <TM4>
F;267-290/Domain: transmembrane #status predicted <TM5>
F;529-552/Domain: transmembrane #status predicted <TM6>
F;562-585/Domain: transmembrane #status predicted <TM7>
F;11,57,324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.1%; Score 8; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGDGGGSC 68
|||||
Db 435 AGDGGGSC 442

RESULT 14
A70934
hypothetical glycine-rich protein Rv0578c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: A70934
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: A70934

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1306 <COG>
A;Cross-references: GB:AL021942; GB:NL123456; NID:G3242298; PIDN:CAA17449.1; PID:G290963;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0578c
C;Superfamily: collagen alpha 1(IV) chain

Query Match 3.1%; Score 8; DB 2; Length 1306;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 VSAGDGG 66
|||||
Db 1191 VSAGDGG 1198

RESULT 15
BXSAL
antibacterial protein 1 - Staphylococcus haemolyticus
N;Alternate names: gonococcal growth inhibitor 1
C;Species: Staphylococcus haemolyticus
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 24-Nov-1999
C;Accession: S00599
R;Watson, D.C.; Yaguchi, M.; Bisailon, J.G.; Beaudet, R.; Morosoli, R.
Biochem. J. 252, 87-93, 1988
A;Title: The amino acid sequence of a gonococcal growth inhibitor from Staphylococcus ha
A;Reference number: S00599; MUID:88339821; PMID:3138972
A;Accession: S00599
A;Molecule type: protein
A;Residues: 1-44 <WAT>
C;Superfamily: Staphylococcus haemolyticus antibacterial protein
C;Keywords: antibacterial; antibiotic; blocked amino end; hemolysis
F;1/Modified site: blocked amino end (Met) (probably formylated) #status experimental

Query Match 2.7%; Score 7; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 AAASVAG 62
|||||
Db 9 AAASVAG 15

Search completed: October 1, 2004, 17:11:20
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:55:17 ; Search time 24 Seconds
(without alignments)

555.415 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256

Sequence: 1 MGKYMKCKGGAAGAAVE.....VRGVPILDAGREFWAPVYSI 256

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 segs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.1	250	1 LINK_PSEPA	P50198 pseudomonas
2	8	3.1	272	1 XPA_MOUSE	Q64267 mus musculus
3	8	3.1	352	1 DMPF_ACICA	Q7wtj2 acinetobact
4	8	3.1	452	1 NIFB_RHOCA	P17434 rhodobact
5	8	3.1	492	1 TMS2_HUMAN	O15393 homo sapien
6	8	3.1	599	1 MX87_YEAST	Q99257 saccharomyc
7	8	3.1	601	1 OAR_DROME	P22270 drosophila
8	7	2.7	44	1 GGI1_STAHA	P11697 staphylococ
9	7	2.7	99	1 NULM_ALBCO	P48928 albinaria c
10	7	2.7	114	1 H33_STYLE	P81197 stylonychia
11	7	2.7	114	1 H36_STYLE	P81200 stylonychia
12	7	2.7	114	1 H37_STYLE	P81201 stylonychia
13	7	2.7	114	1 H38_STYLE	P81196 stylonychia
14	7	2.7	118	1 R18E_SULAC	P39474 sulfobolus
15	7	2.7	124	1 CRCB_PSEPK	Q88ft1 pseudomonas
16	7	2.7	134	1 ATPE_RHORU	P05442 rhodospiril
17	7	2.7	142	1 NCBI_YEAST	P40096 saccharomyc
18	7	2.7	149	1 YCIA_ZYMMO	O66120 zymomonas m
19	7	2.7	152	1 MYG_APTFO	P02199 aptenodytes
20	7	2.7	160	1 Y964_MYCTU	P71546 mycobacteri
21	7	2.7	160	1 Y989_MYCBO	P59978 mycobacteri
22	7	2.7	187	1 PAAD_CAMJE	Q9pbf1 campylobact
23	7	2.7	208	1 COBH_MYCTU	Q10676 mycobacteri
24	7	2.7	220	1 Y085_MYCTU	Q10882 mycobacteri
25	7	2.7	230	1 CUTI_FUSSO	P00590 fusarium so
26	7	2.7	230	1 CUTI_FUSSC	Q99174 fusarium so
27	7	2.7	231	1 CUT3_FUSSO	Q96us9 fusarium so
28	7	2.7	242	1 YK81_STRCO	P45497 streptomyc
29	7	2.7	263	1 GNP2_GIALA	O97440 giardia lam
30	7	2.7	289	1 AROK_METTH	O26896 mechanobact
31	7	2.7	296	1 Y4YQ_RHISN	P55725 rhizobium s
32	7	2.7	327	1 THIO_RHIET	O34292 rhizobium e
33	7	2.7	335	1 CPR4_CAEEL	P43508 caenorhabdi

34 7 2.7 337 1 COBU_PSEDE P29335 pseudomonas
35 7 2.7 338 1 COBT_RHIME Q92p98 rhizobium m
36 7 2.7 342 1 COBT_METSQ Q9x7f4 methylobact
37 7 2.7 344 1 COBT_AGR15 Q8ue71 agrobacteri
38 7 2.7 349 1 MOAA_RHIME Q92pb4 rhizobium m
39 7 2.7 353 1 AOX1_TOBAC Q41224 nicotiana t
40 7 2.7 361 1 IF35_MOUSE Q9dch4 mus musculu
41 7 2.7 371 1 PGLR_PENJA O42824 penicillium
42 7 2.7 385 1 VE2_COPV Q89420 canine oral
43 7 2.7 394 1 ATOB_ECOLI P76461 escherichia
44 7 2.7 398 1 THIL_YEAST P41338 saccharomyc
45 7 2.7 406 1 FBW3_HUMAN Q9ukb7 homo sapien

ALIGNMENTS

RESULT 1
LINK_PSEPA
ID LINK_PSEPA STANDARD; PRT; 250 AA.
AC P50198;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2,5-dichloro-2,5-cyclohexadiene-1,4-diol dehydrogenase (EC 1.1.1.1.-)
DE (2,5-DDOL dehydrogenase).
GN LINK.
OS Pseudomonas paucimobilis (Sphingomonas paucimobilis).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Sphingomonas.
OX NCBI_TaxID=13689;
RN [1]_TaxID=13689;
RP SEQUENCE FROM N.A.
RX STRAIN=UT26;
RX MEDLINE=94252977; PubMed=7515041;
RA Nagata Y., Ohtomo R., Miyauchi K., Fukuda M., Yano K., Takagi M.;
RT "Cloning and sequencing of a 2,5-dichloro-2,5-cyclohexadiene-1,4-diol
RT dehydrogenase gene involved in the degradation of gamma-
RT hexachlorocyclohexane in Pseudomonas paucimobilis.";
RL J. Bacteriol. 176:3117-3125(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B90;
RC MEDLINE=22338262; PubMed=12450824;
RX Kumari R., Subudhi S., Suar M., Dhinra G., Raina V., Dogra C.,
RA Lal S., van der Meer J.R., Holliger C., Lal R.;
RT "Cloning and characterization of lin genes responsible for the
RT degradation of hexachlorocyclohexane isomers by Sphingomonas
RT paucimobilis strain B90.";
RL Appl. Environ. Microbiol. 68:6021-6028(2002).
CC -!- FUNCTION: Catalyzes the degradation of 2,5-dichloro-2,5-
CC cyclohexadiene-1,4-diol (2,5-DDOL) into 2,5-dichlorohydroquinone
CC (2,5-DCHQ). Link appears not to be involved in gamma-HCH
CC degradation pathway.
CC -!- CATALYTIC ACTIVITY: 2,5-DDOL + NAD(+) = 2,5-DCHQ + NADH.
CC -!- SIMILARITY: belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D23722; BAA04939.1; -;
CC EMBL: AY150579; RAN64237.1; -;
CC HSSP: P19992; IHDC.
CC InterPro: IPR002198; ADH short.
CC Pfam: PF00106; adh short; 1.
CC PRINTS: PR00080; SDRFAMILY.
CC PROSITE: PS00061; ADH_SHORT; 1.
CC

```
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NP_BIND 9 34 NAD (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT VARIANT 80 81 DG -> EA (IN STRAIN B90).
SQ SEQUENCE 250 AA; 25492 MW; 8C52703FF76382CF CRC64;

Query Match 3.1%; Score 8; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 7.1; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 222 AAQAKRFA 229
Db 21 AAQAKRFA 28
|||||

RESULT 2
XPB_MOUSE
ID_XPB_MOUSE STANDARD; PRT; 272 AA.
AC O64267;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-repair protein complementing XP-A cells homolog (Xeroderma
pigmentosum group A complementing protein homolog).
GN XPA OR XPAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX MEDLINE=94173654; PubMed=8127648;
RA van Oostrom C.T.N., de Vries A., Verbeek S.J., van Kreijl C.F.,
RA van Sreeg H.;
RT "Cloning and characterization of the mouse XPAC gene.";
RL Nucleic Acids Res. 22:11-14(1994).
CC -!- FUNCTION: Involved in DNA excision repair. Initiates repair by
CC binding to damaged sites with various affinities, depending on the
CC photoproduct and the transcriptional state of the region (By
CC similarity).
CC -!- SUBUNIT: Interacts with XAB1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the XPA family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

DR EMBL; X74345; CAA52392.1; -.
DR EMBL; X74346; CAA52392.1; JOINED.
DR EMBL; X74347; CAA52392.1; JOINED.
DR EMBL; X74348; CAA52392.1; JOINED.
DR EMBL; X74349; CAA52392.1; JOINED.
DR EMBL; X74350; CAA52392.1; JOINED.
DR EMBL; X74351; CAA52393.1; -.
DR PIR; S41498; S41498.
DR MGD; MGI:99135; Xpa.
DR GO; GO:006289; P:nucleotide-excision repair; IMP.
DR GO; GO:0006979; P:response to oxidative stress; IMP.
DR InterPro; IPR000465; XPA protein.
DR Pfam; PF05181; XPA_C; 1.
DR Pfam; PF01286; XPA_N; 1.
DR TIGRFAMs; TIGR00598; rad14; 1.
DR PROSITE; PS00752; XPA_1; 1.
DR PROSITE; PS00753; XPA_2; 1.
KW DNA repair; DNA-binding; Zinc-finger; Nuclear protein.
FT DOMAIN 28 49 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 78 82 POLY-GLU.

FT ZN_FING 104 128
SQ SEQUENCE 272 AA; 31399 MW; 9061C952817E8DCB CRC64;

Query Match 3.1%; Score 8; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 7.6; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 34 AAATGGVA 41
Db 53 AAATGGVA 60
|||||

RESULT 3
DMP ACICA
ID_DMP ACICA STANDARD; PRT; 352 AA.
AC Q7WUJ2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phenol hydroxylase P5 protein (EC 1.14.13.7) (Phenol 2-monooxygenase
P5 component).
GN MHP.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PHEA-2;
RX MEDLINE=22617848; PubMed=12732969;
RA Xu Y., Chen M., Zhang W., Lin M.;
RT "Genetic organization of genes encoding phenol hydroxylase, benzoate
1,2-dioxygenase alpha subunit and its regulatory proteins in
Acinetobacter calcoaceticus PHEA-2.";
RL Curr. Microbiol. 46:235-240(2003).
RN [2]
RP SEQUENCE OF 1-25, AND INDUCTION.
RC STRAIN=69-V;
RA Benndorf D., Davidson I., Babel W.;
RT "Growth on phenol at chemotress levels enhances induction of the
phenol degradation pathway in Acinetobacter calcoaceticus.";
RL Microbiology 0:0-0(2003).
CC -!- FUNCTION: Catabolizes phenol, and some of its methylated
CC derivatives. P5 is required for growth on phenol, and for in vitro
CC phenol hydroxylase activity (By similarity).
CC -!- FUNCTION: Probable electron transfer from NADPH, via FAD and the
CC 2Fe-2S center, to the oxygenase activity site of the enzyme (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Phenol + NADPH + O(2) = catechol + NADP(+) +
CC H(2)O.
CC -!- COFACTOR: P5 contains a FAD cofactor, and a ferredoxin-type
CC [2Fe-2S] center (By similarity).
CC -!- PATHWAY: Phenol biodegradation; first step.
CC -!- SUBUNIT: The multicomponent enzyme phenol hydroxylase is formed by
CC P0, P1, P2, P3, P4 and P5 polypeptides (By similarity).
CC -!- INDUCTION: By phenol.
CC -!- SIMILARITY: In the N-terminal section; belongs to the 2Fe2S plant-
CC type ferredoxin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

DR EMBL; AJ564846; CAD92316.1; -.
DR InterPro; IPR006058; 2Fe2S fd BS.
DR InterPro; IPR008333; FAD binding_6.
DR InterPro; IPR01041; Ferredoxin.
DR InterPro; IPR010205; NqrF.
DR InterPro; IPR001433; Oxred_FAD/NAD(P).
```

DR Pfam; PF00970; FAD binding_6; 1.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF00175; NAD binding_1; 1.
DR TIGRfams; TIGR02008; fdx_plant; 1.
DR TIGRfams; TIGR01941; ngrF; 1.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase;
KW Monooxygenase; Flavoprotein; FAD; NADP;
KW Metal-binding; Iron-sulfur; Iron; 2Fe-2S;
KW Electron transport.
FT INIT_MET 0
FT METAL 36 36 IRON-SULFUR (2FE-2S) (BY
SIMILARITY).
FT METAL 41 41 IRON-SULFUR (2FE-2S) (BY
SIMILARITY).
FT METAL 44 44 IRON-SULFUR (2FE-2S) (BY
SIMILARITY).
FT METAL 76 76 IRON-SULFUR (2FE-2S) (BY
SIMILARITY).
SQ SEQUENCE 352 AA; 38906 MW; D15C8632F5620B7D CRC64;
Query Match 3.1%; Score 8; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 AADGAAGQ 104
DB 336 AADGAAGQ 343
RESULT 4
NIFB RHOC
ID NIFB RHOC STANDARD; PRT; 452 AA.
AC P17434;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE FeMo cofactor biosynthesis protein nifb.
GN NIFB.
OS Rhodospirillum rubrum (Rhodospirillum rubrum)
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Rhodospirillum.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88232432; PubMed=2836706;
RA Masepohl P., Klipp W., Puchler A.;
RT "Genetic characterization and sequence analysis of the duplicated
nifA/nifB gene region of Rhodospirillum rubrum";
RL Mol. Gen. Genet. 212:27-37(1988).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE SYNTHESIS OF THE FE-MO
COFACTOR.
CC -!- SIMILARITY: BELONGS TO THE MOA / NIFB / PQOE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X07567; CAA30450.1; --
DR PIR; S03829; S03829.
DR InterPro; IPR003731; DUF153.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR000385; MoaA_NifB_PqqE.
DR InterPro; IPR005980; NifB.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF02579; Nitro_FeMo-Co; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRfams; TIGR01290; nifB; 1.

DR PROSITE; PS01305; MOA_NIFB_PQOE; 1.
KW Nitrogen fixation; Iron-sulfur.
FT METAL 34 34 IRON-SULFUR (POTENTIAL).
FT METAL 38 38 IRON-SULFUR (POTENTIAL).
FT METAL 41 41 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 452 AA; 49445 MW; EB6A7FEE31DDAD7D CRC64;
Query Match 3.1%; Score 8; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 218 AAEAAQA 225
DB 316 AAEAAQA 323
RESULT 5
TMS2 HUMAN
ID TMS2 HUMAN STANDARD; PRT; 492 AA.
AC Q15393; Q9BXK1;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 2 precursor (EC 3.4.21.-).
GN TMPSR2 OR PRSS10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97468144; PubMed=9325052;
RA Paoloni-Giacobino A., Chen H., Peitsch M.C., Rossier C.,
Antonarakis S.E.;
RT "Cloning of the TMPSR2 gene, which encodes a novel serine protease
with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3";
RL Genomics 44:309-320(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21309069; PubMed=11414763;
RA Teng D.H., Chen Y., Lian L., Ha P.C., Tavtigian S.V., Wong A.K.;
RT "Mutation analyses of 268 candidate genes in human tumor cell lines";
RL Genomics 74:352-364(2001).
RN [3]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RX MEDLINE=21139112; PubMed=11245484;
RA Afar D.B.H., Vivanco I., Hubert R.S., Kuo J., Chen E., Saffran D.C.,
Raitano A.B., Jakobovits A.;
RT "Catalytic cleavage of the androgen-regulated TMPSR2 protease results
in its secretion by prostate and prostate cancer epithelia";
RL Cancer Res. 61:1686-1692(2001).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=21104370; PubMed=11169526;
RA Vaarala M.H., Porvari K.S., Kellokumpu S., Kyllonen A.P., Viikio P.T.;
RT "Expression of transmembrane serine protease TMPSR2 in mouse and
human tissues";
RL J. Pathol. 193:134-140(2001).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Activated by
cleavage and secreted.
CC -!- TISSUE SPECIFICITY: Expressed strongly in small intestine. Also
expressed in prostate, colon, stomach and salivary gland.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

CC --!- SIMILARITY: Belongs to the NXF family.
 CC --!- SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
 CC --!- SIMILARITY: Contains 4 LRR domain.
 CC --!- SIMILARITY: Contains 1 UBA-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X96770; CAA65552.1; -;
 CC EMBL; 273525; CAA97875.1; -;
 CC PIR; S65180; S65180.
 CC SGD; S0006090; MEX67.
 CC GO; GO:0005737; C:cytoplasm; IDA.
 CC GO; GO:0005643; C:nuclear pore; IDA.
 CC GO; GO:0005515; F:protein binding; IPI.
 CC GO; GO:0003723; F:RNA binding; IDA.
 CC GO; GO:0017056; F:structural constituent of nuclear pore; IDA.
 CC GO; GO:0006406; F:mRNA-nucleus export; IMP.
 CC InterPro; IPR002075; NTF2.
 CC Pfam; PF03943; TAP C; 1 -
 CC PROSITE; PS50177; NTF2_DOMAIN; 1.
 CC Transprot; mRNA transport; Nuclear protein; Repeat;
 CC Leucine-rich repeat.
 CC REPEAT 163 188 LRR 1.
 CC REPEAT 189 214 LRR 2.
 CC REPEAT 215 238 LRR 3.
 CC REPEAT 239 264 LRR 4.
 CC DOMAIN 280 467 NTF2.
 CC DOMAIN 547 587 UBA-LIKE (BY SIMILARITY).
 CC MUTAGEN 400 400 H->Y: IMPAIRS ASSOCIATION WITH THE
 CC NUCLEAR PORES AND INTERACTION WITH MTR2.
 CC SQ SEQUENCE 599 AA; 67351 MW; 6C555BEC6FA0B40 CRC64;
 CC
 CC Query Match 3.1%; Score 8; DB 1; Length 599;
 CC Best Local Similarity 100.0%; Pred. No. 15;
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 81 PPQPPSV 88
 CC |||||
 CC Db 506 PPQPPSV 513
 CC
 CC RESULT 7
 CC OAR_DROME STANDARD; PRT; 601 AA.
 CC AC P22270; Q9VNW3;
 CC DT 01-AUG-1991 (Rel. 19, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Tyramine/octopamine receptor precursor (Tyr/Oct-Dro).
 CC GN TYRR OR OCR OR TYR OR OCTYR99AB OR CG7485.
 CC OS Drosophila melanogaster (Fruit fly).
 CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC OC Ephydroidea; Drosophilidae; Drosophila.
 CC OX NCBI_TaxID=7227;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=Oregon-R;
 CC RX MEDLINE=90198291; PubMed=2156539;
 CC RA Arakawa S., Gocayne J.D., McCombie W.R., Urquhart D.A., Hall L.M.,
 CC RA Fraser C.M., Venter J.C.;
 CC RT "Cloning, localization, and permanent expression of a Drosophila
 CC RT octopamine receptor."
 CC RL Neuron 4:343-354(1990).
 CC RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R; TISSUE=Head;
 RX MEDLINE=91006061; PubMed=2170118;
 RA Saudou F., Amlaiky N., Plassat J.-L., Borelli E., Hen R.;
 RT "Cloning and characterization of a Drosophila tyramine receptor."
 RL EMBO J. 9:3611-3617(1990).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablics B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC --!- FUNCTION: Receptor for both octopamine and tyramine, invertebrate
 CC neurotransmitters, and neuromodulators. The activity of this
 CC receptor is mediated by G proteins which activate adenylyl
 CC cyclase. The rank order of potency for agonists is tyramine >
 CC octopamine > dopamine > epinephrine > norepinephrine > serotonin >
 CC histamine. For antagonists, the rank order is yohimbine >
 CC chlorpromazine > phentolamine > mianserin > cyproheptadine >
 CC dihydroergotamine > clonidine > synephrine.
 CC --!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC --!- TISSUE SPECIFICITY: Preferentially expressed in Drosophila heads.
 CC --!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M60789; AAA28731.1; -;
 CC EMBL; X54794; CAA38565.1; -;
 CC EMBL; AE003596; AAF51802.1; -;
 CC PIR; S12004; S12004.
 CC HSSP; P29274; 1MMH.

```
DR FlyBase: FBgn0004514; Tytr.  
DR GO: GO:0007608; Prolactation; IMP.  
DR InterPro: IPR000276; GPCR Rhodpsn.  
DR Pfam: PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR RHODOPSIN.  
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.  
DR PROSITE; PS0262; G PROTEIN RECEPTOR FL 2; 1.  
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.  
KW SIGNAL 1 26  
FT CHAIN 27 601 TYRAMINE/OCTOPAMINE RECEPTOR.  
FT DOMAIN 27 112 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 113 135 1 (POTENTIAL).  
FT DOMAIN 136 145 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 146 167 2 (POTENTIAL).  
FT DOMAIN 168 184 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 185 205 3 (POTENTIAL).  
FT DOMAIN 206 225 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 226 248 4 (POTENTIAL).  
FT DOMAIN 249 273 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 274 295 5 (POTENTIAL).  
FT DOMAIN 296 530 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 531 552 6 (POTENTIAL).  
FT DOMAIN 553 566 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 567 588 7 (POTENTIAL).  
FT DOMAIN 589 601 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 182 261 BY SIMILARITY.  
FT CONFLICT 34 34 S -> N (IN REF. 1).  
SQ SEQUENCE 601 AA; 64674 MW; 7E7581A11674B4C9 CRC64;  
  
Query Match 3.1%; Score 8; DB 1; Length 601;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 61 AGGDDGGSC 68  
DB 435 AGGDDGGSC 442  
  
RESULT 8  
GGII STAAH STANDARD; PRT; 44 AA.  
AC P11697;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Antibacterial protein 1 (Gonococcal growth inhibitor 1).  
OS Staphylococcus haemolyticus.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI_TaxID=1283;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=88399821; PubMed=3138972;  
RA Watson D.C., Yaguchi M., Bisailon J.G., Beaudet R., Morosoli R.;  
RT "The amino acid sequence of a gonococcal growth inhibitor from  
RT Staphylococcus haemolyticus."  
RL Biochem. J. 252:87-93(1989).  
CC -!- DOMAIN: THIS PROTEIN IS POSSIBLY THE SIGNAL SEQUENCE OF A SECRETED  
CC -!- OR MEMBRANE ASSOCIATED PROTEIN.  
CC -!- SIMILARITY: HIGH, TO ANTIBACTERIAL PROTEINS 2 AND 3.  
CC PIR; S00599; BXSAL.  
DR InterPro: IPR008846; Staph.baemo.  
DR Pfam; PF05480; Staph.baemo; 1.  
KW Antibiotic; Formylation.  
FT MOD RES 1 1 FORMYLATION (POTENTIAL).  
SQ SEQUENCE 44 AA; 4523 MW; FABI702DBA3A3238 CRC64;  
  
Query Match 2.7%; Score 7; DB 1; Length 44;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 56 AAAYSAG 62
```

```
Db 9 AAAYSAG 15  
|||||  
RESULT 9  
NULM ALBCO STANDARD; PRT; 99 AA.  
ID NULM ALBCO  
AC P48928;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).  
GN ND4L.  
OS Albinaria coerulea (Land snail).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylomatophora;  
OC Sigmurethra; Clausilloidea; Clausiliidae; Aloiinae; Albinaria.  
OX NCBI_TaxID=42349;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96120351; PubMed=7498775;  
RA Hatzoglou E., Rodakis G.C., Lecanidou R.;  
RT "Complete sequence and gene organization of the mitochondrial genome  
RT of the land snail Albinaria coerulea."  
RL Genetics 140:1353-1366(1995).  
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; X83390; CAA58298.1; -.  
DR PIR; S59145; S59145.  
DR InterPro: IPR001133; Oxidored 4L.  
DR Pfam; PF00420; oxidored g2; 1.  
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
SQ SEQUENCE 99 AA; 10909 MW; CB65A360B646AAES CRC64;  
  
Query Match 2.7%; Score 7; DB 1; Length 99;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 217 FAAAAEA 223  
DB 63 FAAAAEA 69  
|||||  
RESULT 10  
H33_STYLE STANDARD; PRT; 114 AA.  
ID H33_STYLE  
AC P81197;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Histone H3-3 (Fragment).  
GN H3-3.  
OS Stylonychia lemnae.  
OC Eukaryota; Alveolata; Ciliophora; Spirotriches; Stichotrichia;  
OC Stichotrichida; Oxytrichidae; Stylonychia.  
OX NCBI_TaxID=5949;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bernhard D.;  
RT "Several highly divergent histone H3 genes in the hypotrich ciliate  
RT Stylonychia lemnae."  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Histone H3, along with histone H4, plays a central role  
CC in nucleosome formation.  
CC -!- SUBUNIT: The nucleosome is an octamer containing two molecules
```

```
CC      each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC      bp of DNA.
CC      -!- SIMILARITY: Belongs to the histone H3 family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; Y16629; CAA76332.1; -
CC      InterPro; IPR007124; Hist_TAF.
CC      InterPro; IPR007125; Histone_core_D.
CC      Pfam; PF00125; histone; 1.
CC      PRINTS; PR00622; HISTONEH3.
CC      SMART; SM00428; H3; 1.
CC      PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
CC      PROSITE; PS00959; HISTONE_H3_2; 1.
CC      Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
CC      FT NON_TER 1 114
CC      SEQUENCE 114 AA; 12721 MW; 36A9B6E8F51D01FC CRC64;
CC
CC      Query Match 2.7%; Score 7; DB 1; Length 114;
CC      Best Local Similarity 100.0%; Pred.No. 31;
CC      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC      QY 34 AAATGGV 40
CC      DB 22 AAATGGV 28
CC      |||||
CC
CC      RESULT 11
CC      H37_STYLE
CC      ID H37_STYLE STANDARD; PRT; 114 AA.
CC      AC P81201;
CC      DT 15-JUL-1998 (Rel. 36, Created)
CC      DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC      DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC      DE Histone H3-6 (Fragment).
CC      GN H3-6.
CC      OS Stylonychia lemnae.
CC      OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
CC      OC Stichotrichida; Oxytrichidae; Stylonychia.
CC      OX NCBI_TaxID=5949;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RA Bernhard D.;
CC      RT "Several highly divergent histone H3 genes in the hypotrich ciliate
CC      Stylonychia lemnae."
CC      RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: Histone H3, along with histone H4, plays a central role
CC      in nucleosome formation.
CC      -!- SUBUNIT: The nucleosome is an octamer containing two molecules
CC      each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC      bp of DNA.
CC      -!- SIMILARITY: Belongs to the histone H3 family.
CC      -----
CC      EMBL; Y16632; CAA76335.1; -
CC      InterPro; IPR007124; Hist_TAF.
CC      InterPro; IPR007125; Histone_core_D.
CC      Pfam; PF00125; histone; 1.
CC      PRINTS; PR00622; HISTONEH3.
CC      SMART; SM00428; H3; 1.
CC      PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
CC      PROSITE; PS00959; HISTONE_H3_2; 1.
CC      Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
CC      FT NON_TER 1 114
CC      SEQUENCE 114 AA; 12689 MW; 36A9B6EA7F1D01FC CRC64;
CC
CC      Query Match 2.7%; Score 7; DB 1; Length 114;
CC      Best Local Similarity 100.0%; Pred.No. 31;
CC      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC      QY 34 AAATGGV 40
CC      DB 22 AAATGGV 28
CC      |||||
CC
CC      RESULT 11
CC      H36_STYLE
CC      ID H36_STYLE STANDARD; PRT; 114 AA.
CC      AC P81200;
CC      DT 15-JUL-1998 (Rel. 36, Created)
CC      DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC      DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC      DE Histone H3-6 (Fragment).
CC      GN H3-6.
CC      OS Stylonychia lemnae.
CC      OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
CC      OC Stichotrichida; Oxytrichidae; Stylonychia.
CC      OX NCBI_TaxID=5949;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RA Bernhard D.;
CC      RT "Several highly divergent histone H3 genes in the hypotrich ciliate
CC      Stylonychia lemnae."
CC      RL Submitted (FEB-1998) to Swiss-Prot.
CC      -!- FUNCTION: Histone H3, along with histone H4, plays a central role
CC      in nucleosome formation.
CC      -!- SUBUNIT: The nucleosome is an octamer containing two molecules
CC      each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC      bp of DNA.
CC      -!- SIMILARITY: Belongs to the histone H3 family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; Y16632; CAA76335.1; -
CC      InterPro; IPR007124; Hist_TAF.
CC      InterPro; IPR007125; Histone_core_D.
CC      Pfam; PF00125; histone; 1.
CC      PRINTS; PR00622; HISTONEH3.
CC      SMART; SM00428; H3; 1.
CC      PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
CC      PROSITE; PS00959; HISTONE_H3_2; 1.
CC      Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
CC      FT NON_TER 1 114
CC      SEQUENCE 114 AA; 12689 MW; 36A9B6EA7F1D01FC CRC64;
CC
CC      Query Match 2.7%; Score 7; DB 1; Length 114;
CC      Best Local Similarity 100.0%; Pred.No. 31;
CC      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC      QY 34 AAATGGV 40
CC      DB 22 AAATGGV 28
CC      |||||
```

```
DR      Pfam; PF00125; histone; 1.
DR      PRINTS; PR00622; HISTONEH3.
DR      SMART; SM00428; H3; 1.
DR      PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
DR      PROSITE; PS00959; HISTONE_H3_2; 1.
DR      Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
DR      FT NON_TER 1 114
DR      SEQUENCE 114 AA; 12794 MW; 0E411BED51DCB06 CRC64;
DR
DR      Query Match 2.7%; Score 7; DB 1; Length 114;
DR      Best Local Similarity 100.0%; Pred.No. 31;
DR      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR
DR      QY 34 AAATGGV 40
DR      DB 22 AAATGGV 28
DR      |||||
DR
DR      RESULT 12
DR      H37_STYLE
DR      ID H37_STYLE STANDARD; PRT; 114 AA.
DR      AC P81201;
DR      DT 15-JUL-1998 (Rel. 36, Created)
DR      DT 15-JUL-1998 (Rel. 36, Last sequence update)
DR      DT 28-FEB-2003 (Rel. 41, Last annotation update)
DR      DE Histone H3-7 (Fragment).
DR      GN H3-7.
DR      OS Stylonychia lemnae.
DR      OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
DR      OC Stichotrichida; Oxytrichidae; Stylonychia.
DR      OX NCBI_TaxID=5949;
DR      RN [1]
DR      RP SEQUENCE FROM N.A.
DR      RA Bernhard D.;
DR      RT "Several highly divergent histone H3 genes in the hypotrich ciliate
DR      Stylonychia lemnae."
DR      RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR      -!- FUNCTION: Histone H3, along with histone H4, plays a central role
DR      in nucleosome formation.
DR      -!- SUBUNIT: The nucleosome is an octamer containing two molecules
DR      each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
DR      bp of DNA.
DR      -!- SIMILARITY: Belongs to the histone H3 family.
DR      -----
DR      This SWISS-PROT entry is copyright. It is produced through a collaboration
DR      between the Swiss Institute of Bioinformatics and the EMBL outstation -
DR      the European Bioinformatics Institute. There are no restrictions on its
DR      use by non-profit institutions as long as its content is in no way
DR      modified and this statement is not removed. Usage by and for commercial
DR      entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR      or send an email to license@isb-sib.ch).
DR      -----
DR      EMBL; Y16633; CAA76336.1; -
DR      InterPro; IPR007124; Hist_TAF.
DR      InterPro; IPR007125; Histone_core_D.
DR      InterPro; IPR00164; Histone_H3.
DR      Pfam; PF00125; histone; 1.
DR      PRINTS; PR00622; HISTONEH3.
DR      SMART; SM00428; H3; 1.
DR      PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
DR      PROSITE; PS00959; HISTONE_H3_2; 1.
DR      Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
DR      FT NON_TER 1 114
DR      SEQUENCE 114 AA; 12689 MW; 36A9B6EA7F1D01FC CRC64;
DR
DR      Query Match 2.7%; Score 7; DB 1; Length 114;
DR      Best Local Similarity 100.0%; Pred.No. 31;
DR      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR
DR      QY 34 AAATGGV 40
DR      DB 22 AAATGGV 40
DR      |||||
```

```
Db 22 AAATGGV 28

RESULT 13
H39_STYLE STANDARD; PRT; 114 AA.
AC P81196;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H3-2 (Fragment)
GN H3-2.
OS Stylonychia lemnae.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
OC Stichotrichida; Oxytrichidae; Stylonychia.
OX NCBI_TaxID=5949;
RN [1]
RP SEQUENCE FROM N.A.
RA Bernhard D.;
RT "Several highly divergent histone H3 genes in the hypotrich ciliate Stylonychia lemnae.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Histone H3, along with histone H4, plays a central role in nucleosome formation.
CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules each of H2A, H2B, H3 and H4. The octamer wraps approximately 146 bp of DNA.
CC -1- SIMILARITY: Belongs to the histone H3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL; AB016788; AAN69595.1; -.
DR TIGR; PP4001; -.
DR HAMAP; MF_00454; -. 1.
DR InterPro; IPR003691; Camphor_CrcB.
DR Pfam; PF02537; CRCB; 1.
KW Transmembrane; Complete proteome.
FT TRANSMEM 4 21 Potential.
FT TRANSMEM 34 56 Potential.
FT TRANSMEM 66 85 Potential.
FT TRANSMEM 97 119 Potential.
SQ SEQUENCE 124 AA; 12973 MW; E1A38ED15D0D55A6 CRC64;

Query Match      2.7%; Score 7; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 AAVSAGG 63
Db      6 AAVSAGG 12
      |||||

```

Search completed: October 1, 2004, 17:08:32
 Job time : 27 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 17:02:22 ; Search time 116 Seconds
(without alignments)
696.316 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256

Sequence: 1 MCKYMKRCGAGAEVAAYE.....VRGVPLDAGGRFEPVVSII 256

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- SPTREMBL 25:*
- 1: sp archaea:*
 - 2: sp bacteria:*
 - 3: sp fungi:*
 - 4: sp human:*
 - 5: sp invertebrate:*
 - 6: sp mammal:*
 - 7: sp mhc:*
 - 8: sp organelle:*
 - 9: sp phase:*
 - 10: sp plant:*
 - 11: sp rodent:*
 - 12: sp virus:*
 - 13: sp vertebrate:*
 - 14: sp unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	3.9	277	17 Q9HP13	Q9hp13 halobacteri
2	9	3.5	591	16 Q8GP35	Q8gp35 bordetella
3	9	3.5	664	16 Q9KRL3	Q9rkl3 streptomyce
4	8	3.1	55	9 Q853R3	Q853r3 mycobacteri
5	8	3.1	173	10 Q9M2I3	Q9m2i3 arabidopsis
6	8	3.1	180	10 Q9H4B9	Q9h4b9 oryza sativ
7	8	3.1	181	10 Q9M2I6	Q9m2i6 arabidopsis
8	8	3.1	191	16 Q89S92	Q89s92 bradyrhizob
9	8	3.1	201	2 Q8GCV1	Q8gcv1 heliobacill
10	8	3.1	206	16 Q8XWK7	Q8xwk7 ralstonia s
11	8	3.1	215	10 Q9M3U5	Q9m3u5 nicotiana t
12	8	3.1	215	10 Q49111	Q49111 nicotiana t
13	8	3.1	221	16 Q98ET3	Q98et3 rhizobium l
14	8	3.1	238	16 Q9FYA5	Q8fyas brucella su
15	8	3.1	248	16 Q89P10	Q89p10 bradyrhizob
16	8	3.1	250	16 Q7U6X9	Q7u6x9 synechococc

17	8	3.1	273	11 Q9CVA0	Q9cva0 mus musculu
18	8	3.1	278	2 Q937M8	Q937m8 photorhabdu
19	8	3.1	279	11 Q8KX7	Q8kx7 mus musculu
20	8	3.1	281	2 Q8KU07	Q8ku07 xenophilus
21	8	3.1	295	10 Q8GY40	Q8gy40 arabidopsis
22	8	3.1	320	16 Q9EMJ3	Q9ejm3 streptomyce
23	8	3.1	331	16 Q7WDV3	Q7wdv3 bordetella
24	8	3.1	331	16 Q7W2V3	Q7w2v3 bordetella
25	8	3.1	334	16 Q865I3	Q865i3 streptomyce
26	8	3.1	336	16 Q9KZF0	Q9kzf0 streptomyce
27	8	3.1	353	2 Q7WTJ2	Q7wtj2 acinetobact
28	8	3.1	396	12 Q8VA04	Q8va04 apple stem
29	8	3.1	401	8 Q9GCV9	Q9gcv9 scherffelia
30	8	3.1	404	16 Q862G3	Q862g3 mycobacteri
31	8	3.1	404	16 Q7TWH6	Q7twh6 mycobacteri
32	8	3.1	410	10 Q7XX17	Q7xx17 oryza sativ
33	8	3.1	418	16 Q8XZK5	Q8xxk5 ralstonia s
34	8	3.1	427	16 Q92BJ2	Q92bj2 listeria in
35	8	3.1	427	16 Q8Y707	Q8y707 listeria mo
36	8	3.1	469	16 Q88AX9	Q88ax9 pseudomonas
37	8	3.1	491	16 Q82B50	Q82b50 streptomyce
38	8	3.1	492	4 Q96T73	Q96t73 homo sapien
39	8	3.1	493	16 Q7UI25	Q7ui25 rhodospirell
40	8	3.1	504	16 Q82T80	Q82t80 nitrosomona
41	8	3.1	548	10 Q9LJF7	Q9ljf7 arabidopsis
42	8	3.1	601	5 Q95YF4	Q95yf4 drosophila
43	8	3.1	618	3 Q86ZN5	Q86zn5 podospora a
44	8	3.1	653	10 Q8H8N3	Q8h8n3 oryza sativ
45	8	3.1	667	5 Q9VFJ6	Q9vfv6 drosophila

ALIGNMENTS

RESULT 1

Q9HP13 PRELIMINARY; PRT; 277 AA.
ID Q9HP13
AC Q9HP13;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Cytochrome aa3 controlling protein.
GN CYP OR VNG1623G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J., Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A., Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W., Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.; Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL "Genome sequence of Halobacterium species NRC-1"; EMBL; AE005072; RAG1984.1; -;
DR PIR; H84314; H84314.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006461; P:protein complex assembly; IEA.
DR InterPro; IPR003780; COX15_CtaA.
DR Pfam; PF02628; COX15_CtaA; 1.
KW Complete proteome.
SQ SEQUENCE 277 AA; 29053 MW; B053123766E274B1 CRC64;

Query Match 3.9%; Score 10; DB 17; Length 277;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 R5AATGGA 41
 |||||
 Db 216 R5AATGGA 225

RESULT 2

Q8GP35 PRELIMINARY; PRT; 591 AA.
 ID Q8GP35
 AC Q8GP35;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Multicopper oxidase (Copper resistance protein).
 GN COFA OR BP3315.
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2224525; PubMed=12354238;
 RA Huston W.N., Jennings M.P., McEwan A.G.;
 RT "The multicopper oxidase of *Pseudomonas aeruginosa* is a ferroxidase
 with a central role in iron acquisition.";
 RL Mol. Microbiol. 45:1741-1750 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
Bordetella parapertussis and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40 (2003).
 DR EMBL; AF455754; AAN52533.1; -;
 DR EMBL; BX640421; CAE43580.1; -;
 DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.
 DR GO; GO:0005488; F:binding; IEA.
 DR GO; GO:0005507; F:copper ion binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR006376; COPA.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR008972; Cupredoxin.
 DR InterPro; IPR002355; Cu ox. copper BS.
 DR InterPro; IPR001993; Mitoch carrier.
 DR InterPro; IPR006311; Tat.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR TIGRfams; TIGR01480; copper res A; 1.
 DR TIGRfams; TIGR01409; TAT signal_seq; 1.
 DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
 DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 KW Complete proteome.
 SQ SEQUENCE 591 AA; 65643 MW; BFA5587C557DE393 CRC64;

Query Match 3.5%; Score 9; DB 16; Length 591;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 GAAGQGGAA 108
 |||||
 Db 384 GAAGQGGAA 392

RESULT 3

Q9RKL3 PRELIMINARY; PRT; 664 AA.
 ID Q9RKL3
 AC Q9RKL3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative integral membrane protein.
 GN SC04069 OR SCD25.05
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete *Streptomyces*
coelicolor A3(2).";
 RL Nature 417:141-147 (2002).
 DR EMBL; AL939118; CAB56349.1; -;
 KW Complete proteome.
 SQ SEQUENCE 664 AA; 69158 MW; C3F83EC120526F2A CRC64;
 Query Match 3.5%; Score 9; DB 16; Length 664;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGLS 114
 |||||
 Db 527 GAALAAAGLS 535

RESULT 4
 Q853R3 PRELIMINARY; PRT; 55 AA.
 ID Q853R3
 AC Q853R3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Gp14.
 OS Mycobacteriophage Bx1.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
 OX NCBI_TaxID=205877;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22592660; PubMed=12705866;
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
 RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
 RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
 RA Hatfull G.F.;
 RT "Origins of highly mosaic mycobacteriophage genomes.";
 RL Cell 113:171-182 (2003).
 DR EMBL; AY129337; AAN16674.1; -;
 SQ SEQUENCE 55 AA; 5697 MW; E4EBAC5A9B3AC3EB CRC64;

Query Match 3.1%; Score 8; DB 9; Length 55;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 AALAAAGLS 114
 |||||
 Db 48 AALAAAGLS 55


```
Db 14 AATPAEEL 21

RESULT 5
Q9M213 ID Q9M213 PRELIMINARY; PRT; 173 AA.
AC Q9M213;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F9D24.240.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137081; CAB68171.1; -.
DR PIR; T45993; T45993.
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 19724 MW; ACD2BF15B48B0C2D CRC64;

Query Match 3.1%; Score 8; DB 10; Length 173;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
Db 148 LSDLES DL 155

RESULT 6
Q8H4B9 ID Q8H4B9 PRELIMINARY; PRT; 180 AA.
AC Q8H4B9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative translational inhibitor protein.
GN P0048D08.27.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
RT clone:P0048D08";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004269; BAC20708.1; -.
DR InterPro; IPR006175; Endoribon LPSP.
DR InterPro; IPR006056; YjGF-like.
DR Pfam; PF01042; ribonuc L-PSI.
DR TIGRfam; TIGR00004; TIGR00004; 1.
DR PROSITE; PS01094; UPF0076; 1.
SQ SEQUENCE 180 AA; 18818 MW; BCE6915086D91B44 CRC64;

Query Match 3.1%; Score 8; DB 10; Length 180;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 AATPAEEL 205
Db 198 LSDLES DL 205

RESULT 7
Q9M216 ID Q9M216 PRELIMINARY; PRT; 181 AA.
AC Q9M216;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F9D24.210.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137081; CAB68168.1; -.
DR PIR; T45990; T45990.
KW Hypothetical protein.
SQ SEQUENCE 181 AA; 20661 MW; F8EA87E08FBA48B0 CRC64;

Query Match 3.1%; Score 8; DB 10; Length 181;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
Db 154 LSDLES DL 161

RESULT 8
Q89S92 ID Q89S92 PRELIMINARY; PRT; 191 AA.
AC Q89S92;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BIL2513 protein.
GN BIL2513.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=USDA 110;
RC MEDLINE=22484998; PubMed=12597275;
EX Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197 (2002).
DR EMBL; AP005944; BAC47778.1; -.
DR InterPro; IPR001173; Glyco trans 2.
DR InterPro; IPR001092; HLH basic.
DR Pfam; PF00535; Glycos transf_2; 1.
DR PROSITE; PS00038; HLH_1; 1.
KW Complete proteome.
SQ SEQUENCE 191 AA; 20410 MW; 3E1A04A0E09A0413 CRC64;

Query Match 3.1%; Score 8; DB 16; Length 191;
```

Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 QGAALAG 112
|||||
DB 64 QGAALAG 71

RESULT 9
Q8GDV1 PRELIMINARY; PRT; 201 AA.
AC Q8GDV1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Cobinamide kinase (EC 2.7.1.1-) (Fragment).
OS Helicobacillus mobilis.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Hellobacteriaceae;
OC Helicobacillus.
OX NCBI_TaxID=28064;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22337798; PubMed=12446909;
RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,
RA Blankenship R.E.;
RT "Whole-genome analysis of photosynthetic prokaryotes.";
RL Science 298:1616-1620(2002).
RW [2]
RP SEQUENCE FROM N.A.
RA Liolios K.G., Chu L., Ostrovskaya O., Mendybaeva N., Koukharenko V.,
RA Gerdes S., Kyrpides N., Overbeek R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY142871; AAN87475.1;
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR InterPro: IPR003203; COBU.
DR Pfam: PF02283; COBU; 1.
KW Kinase; Transferase.
FT NON TER 201
SQ SEQUENCE 201 AA; 22473 MW; 48D6DC53C37A759 CRC64;

Query Match 3.1%; Score 8; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 AAEAGGDH 148
|||||
DB 27 AAEAGGDH 34

RESULT 10
Q8XWK7 PRELIMINARY; PRT; 206 AA.
AC Q8XWK7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Probable transmembrane protein.
GN RSC2467 OR RS01139.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Cholsine N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissbach J., Boucher C.A.;

"Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646070; CAD16174.1; -.
KW Complete proteome.
SQ SEQUENCE 206 AA; 21783 MW; 8D49E45A150B8440 CRC64;

Query Match 3.1%; Score 8; DB 16; Length 206;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 AALAAGLS 114
|||||
DB 5 AALAAGLS 12

RESULT 11
Q9M3U5 PRELIMINARY; PRT; 215 AA.
AC Q9M3U5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE DRPP2 protein.
GN DRPP2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Bassac G., Domergue O., Grignon N., Ferriere N., Escoute J.,
RA Teyssendier de la Serve B.;
RT "Cloning promoters of two developmentally regulated tobacco genes.
RT Study of their spatio-temporal expression pattern.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277898; CAB91552.1; -.
DR InterPro: IPR008469; DREPP.
DR Pfam: PF05558; DREPP; 1.
SQ SEQUENCE 215 AA; 22915 MW; 2AD0E2F7DEF0346E CRC64;

Query Match 3.1%; Score 8; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 PAATPAE 204
|||||
DB 176 PAATPAE 183

RESULT 12
O49911 PRELIMINARY; PRT; 215 AA.
AC O49911;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Plasma membrane polypeptide.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Logan D.C., Domergue O., Teyssendier de la Serve B., Rossignol M.;
RT "A new family of plasma membrane intrinsic polypeptides differentially
RT regulated during plant development.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y08609; CAA69901.1; -.
DR PIR: T03680; T03680.
DR InterPro: IPR008469; DREPP.

```

DR Pfam: PF05558; DREPP; 1.
SQ SEQUENCE 215 AA; 22943 MW; 9DC7E2F7DEE909D8 CRC64;

Query Match
Best Local Similarity 3.1%; Score 8; DB 10; Length 215;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 PAATPAAE 204
DB 176 PAATPAAE 183

RESULT 13
Q98ET3 PRELIMINARY; PRT; 221 AA.
AC Q98ET3;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein mlr4094.
GN MLR4094.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003003; BAB50834.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006402; HAD-SF-IA-v3.
DR InterPro; IPR005834; Hydrolase.
DR Pfam; PF00702; Hydrolase; 1.
DR TIGRfams; TIGR01509; HAD-SF-IA-v3; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 221 AA; 23018 MW; 477A3DC6B96F6F29 CRC64;

Query Match
Best Local Similarity 3.1%; Score 8; DB 16; Length 221;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AAATAAQA 225
DB 169 AAATAAQA 176

RESULT 14
Q8FYA5 PRELIMINARY; PRT; 238 AA.
AC Q8FYA5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thiol:disulfide interchange protein, putative.
GN BR1980.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;

Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
Ra Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014487; AAN30870.1; -.
DR TIGR; BR1980; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006662; ThioRed.
DR InterPro; IPR006663; ThioRedox_dom2.
DR PROSITE; PS00194; THIOREDOXIN; 1.
KW Complete proteome.
SQ SEQUENCE 238 AA; 24862 MW; F9BAB66EC07B7C58 CRC64;

Query Match
Best Local Similarity 3.1%; Score 8; DB 16; Length 238;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGVA 41
DB 79 AAATGGVA 86

RESULT 15
Q89P10 PRELIMINARY; PRT; 248 AA.
AC Q89P10;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN BLR3673.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005948; BAC48938.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Complete proteome.
SQ SEQUENCE 248 AA; 25207 MW; D137219FCA7FC0C6 CRC64;

Query Match
Best Local Similarity 3.1%; Score 8; DB 16; Length 248;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAGL 113
DB 42 GAALAGL 49

Search completed: October 1, 2004, 17:10:35
Job time : 119 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:54:37 ; Search time 125 Seconds
(without alignments)
578.657 Million cell updates/sec

Title: US-09-993-808b-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVPLDAGRFEWAPVVISI 256

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256	100.0	256	6	ABB98757
2	101	39.5	125	3	AB26250 Corn cycl
3	42	16.4	42	6	AB98860 Maize CKI
4	21	8.2	46	6	AB98862 Maize CKI
5	21	8.2	53	3	AB26245 Corn cycl
6	21	8.2	53	3	AB26245 Corn cycl
7	21	8.2	116	5	AB65674 Corn ICK
8	21	8.2	126	3	AB26249 Corn cycl
9	17	6.6	227	6	AB98759 Maize CKI
10	17	6.6	262	5	AB65670 OsICK 2 p
11	10	3.9	10	5	AB65741 Plant ICK
12	10	3.9	226	5	AB65692 Rice OsIC
13	9	3.5	25	5	AB65793 Plant pot
14	9	3.5	37	3	AB27252 Rice cycl
15	9	3.5	37	3	AB27252 Rice cycl
16	9	3.5	324	6	ABP99213 Orthosomy
17	9	3.5	580	6	AB28345 Trypsin m
18	8	3.1	8	5	AB65707 Plant ICK
19	8	3.1	8	5	AB65767 Plant ICK
20	8	3.1	45	3	AB52025 Human sec
21	8	3.1	62	3	AB98379 Partial t
22	8	3.1	62	3	AB36150 Mosquito
23	8	3.1	62	3	AB37950 Trypsin m
24	8	3.1	62	3	AB28345 Trypsin M
25	8	3.1	62	7	ADC35463 Trypsin m

ALIGNMENTS

RESULT 1

ABB98757
ID ABB98757 standard; protein; 256 AA.

XX AC ABB98757;

XX DT 21-FEB-2003 (first entry)

XX DE Maize CKI_B.

XX KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;
crop yield; root size; plant growth; tassel size; ear size;
male sterility; endoreduplication.

XX OS Zea mays.

XX FH Key Location/Qualifiers
Domain 1..7
/note= "Conserved domain"

XX PN WO200281623-A2.

XX PD 17-OCT-2002.

XX PF 06-NOV-2001; 2001WO-US044038.

XX PR 07-NOV-2000; 2000US-0246349P.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

XX DR WPI; 2003-058511/05.

XX DR N-PSDB; ABV74603.

XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
involved in cell cycle regulation, and useful for altering cell cycle
protein content, cell cycle progression, cell number and composition of
plants.

XX Claim 12; Page 65-66; 69pp; English.

XX The present sequence is maize cyclin-dependent kinase inhibitor (CKI),
CKI_B. The coding sequence for this protein (I) is useful for modulating
the activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
barley, oil-seed Brassica and millet. Modulating the activity of CDK,

Add10243 Mosquito
Aam01295 Human tra
Aau69940 Human tra
Abu71831 Prostate
Abb95400 Human tra
Abr54512 Partial H
Adb14347 Human tra
Abb69862 Drosophil
Abm67642 Phototrab
Abu27612 Protein e
Adb80207 Mycobacte
Abb49079 Listeria
Abu32600 Protein e
Aay44406 Human 20P
Aay77726 Human tum
Aay57280 Ovr115 ho
Aay92050 HrPCa6/7
Aab36901 Human TMP
Aam01294 Human tra
Aam01315 F100C am

CC preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain
 XX
 SQ Sequence 256 AA;

Query Match 100.0%; Score 256; DB 6; Length 256;
 Best Local Similarity 100.0%; Pred. No. 1.7e-234;
 Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKYMRCRGAAGAEEVAEVTVGVVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
 DB 1 MGKYMRCRGAAGAEEVAEVTVGVVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60

QY 61 AGDGGSCYTHLSRMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTA 120
 DB 61 AGDGGSCYTHLSRMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTA 120

QY 121 SSVNLGLGGGSGHTCRSYDAEAGGDHVLVDVSAASNSGSGDPRRRTTPSSRAHGEL 180
 DB 121 SSVNLGLGGGSGHTCRSYDAEAGGDHVLVDVSAASNSGSGDPRRRTTPSSRAHGEL 180

QY 181 SDLESDLAGHKTGSLPAATPAALIVPPAHEIQEFPFAAAEAAQAKFPASKYINDFVRGV 240
 DB 181 SDLESDLAGHKTGSLPAATPAALIVPPAHEIQEFPFAAAEAAQAKFPASKYINDFVRGV 240

QY 241 PLDAGGRFEWAPVSI 256
 DB 241 PLDAGGRFEWAPVSI 256

RESULT 2
 AAB26250
 ID AAB26250 standard; protein; 125 AA.

XX
 AC AAB26250;

DT 17-JAN-2001 (first entry)

DE Corn cyclin-dependent kinase inhibitor #3.

XX Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 XX cell growth; herbicide.

XX Zea mays.

XX WO200060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

DR N-PSDB; AAA95281.

XX

PT Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.

XX Claim 10; Fig 1; 58pp; English.

XX The present sequence is the corn cyclin-dependent kinase inhibitor
 CC (CDKI). Its coding sequence was isolated by searching a contig comprising
 CC cDNA from corn roots and ear leaf sheath for sequences similar to those
 CC encoding the CDKI from *Chenopodium rubrum*, *Caenorhabditis elegans* and
 CC *Arabidopsis thaliana*. CDKI is involved in the cell cycle, and may promote
 CC or inhibit cell division and growth. The protein and its coding sequence
 CC are useful in the production of transgenic plants which produce increased
 CC or decreased amounts of the CDKI protein, in the identification of
 CC herbicides, in genetic and physical mapping and in the isolation of the
 CC CDKI gene in other organisms

XX Sequence 125 AA;

Query Match 39.5%; Score 101; DB 3; Length 125;
 Best Local Similarity 100.0%; Pred. No. 1.4e-87;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 ASNSGSGDPRRRTTPSSRAHGEISDLESDLAGHKTGSLPAATPAALIVPPAHEIQE 215
 DB 25 ASNSGSGDPRRRTTPSSRAHGEISDLESDLAGHKTGSLPAATPAALIVPPAHEIQE 84

QY 216 FFAAAEAAQAKFPASKYINDFVRGVPLDAGGRFEWAPVSI 256

DB 85 FFAAAEAAQAKFPASKYINDFVRGVPLDAGGRFEWAPVSI 125

RESULT 3

ABB98860

ID ABB98860 standard; protein; 42 AA.

XX ABB98860;

DT 21-FEB-2003 (first entry)

XX Maize CKI_B C-terminal protein fragment.

XX Maize; cyclin-dependent kinase inhibitor; CKI; plant; corn; crop yield;
 XX root size; plant growth; tassel size; ear size; male sterility;
 XX endoreduplication.

XX Zea mays.

XX WO200281623-A2.

XX 17-OCT-2002.

XX 06-NOV-2001; 2001WO-US044038.

XX 07-NOV-2000; 2000US-0246349P.

XX (PION-) PIONEER HI-BRED INT INC.

XX (ARIZ-) ARIZONA BOARD OF REGENTS.

XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkens BR, Sun Y;

XX WPI; 2003-058511/05.

XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
 XX involved in cell cycle regulation, and useful for altering cell cycle
 XX protein content, cell cycle progression, cell number and composition of
 XX plants.

XX Example 11; Page 46; 69pp; English.

XX The present invention relates to maize cyclin-dependent kinase inhibitor
 CC (CKI), CKI B, CKI C and CKI D. The coding sequences (I) for these
 CC proteins are useful for modulating the activity of CDK in a plant such as
 CC a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton,

CC rice, barley, oil-seed Brassica and millet. Modulating the activity of
 CC CDK, preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain. The present
 CC sequence is a C-terminal protein fragment of CKI B. This sequence
 CC contains a CDK binding region and/or cyclin binding domains
 XX
 SQ Sequence 42 AA;

Query Match 16.4%; Score 42; DB 6; Length 42;
 Best Local Similarity 100.0%; Pred. No. 5.3e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 EFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVWSI 256
 |||||
 Db 1 EFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVWSI 42

RESULT 4
 ABB98862
 ID ABB98862 standard; protein; 46 AA.

XX
 AC ABB98862;

DT 21-FEB-2003 (first entry)

DE Maize CKI_D C-terminal protein fragment.

XX Maize; cyclin-dependent kinase inhibitor; CKI; plant; corn; crop yield;
 KW root size; plant growth; tassel size; ear size; male sterility;
 KW endoreduplication.

XX Zea mays.

PN W0200281623-A2.

PD 17-OCT-2002.

PF 06-NOV-2001; 2001WO-US044038.

PR 07-NOV-2000; 2000US-0246349P.

XX (PTON-) PIONEER HI-BRED INT INC.
 PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

DR WPI; 2003-058511/05.

XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
 PT involved in cell cycle regulation, and useful for altering cell cycle
 PT protein content, cell cycle progression, cell number and composition of
 PT plants.

PS Example 11; Page 46; 69pp; English.

XX The present invention relates to maize cyclin-dependent kinase inhibitor
 CC (CKI), CKI B, CKI C and CKI D. The coding sequences (I) for these
 CC proteins are useful for modulating the activity of CDK in a plant such as
 CC a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton,
 CC rice, barley, oil-seed Brassica and millet. Modulating the activity of

CC CDK, preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain. The present
 CC sequence is a C-terminal protein fragment of CKI D. This sequence
 CC contains a CDK binding region and/or cyclin binding domains
 XX
 SQ Sequence 46 AA;

Query Match 8.2%; Score 21; DB 6; Length 46;
 Best Local Similarity 100.0%; Pred. No. 5.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
 |||||
 Db 16 AKRFASKYNFDFVRGVPLDAG 36

RESULT 5
 AAB26245

ID AAB26245 standard; protein; 53 AA.

XX AAB26245;

DT 17-JAN-2001 (first entry)

DE Corn cyclin-dependent kinase inhibitor #1.

XX Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 KW cell growth; herbicide.

XX Zea mays.

PN W0200060087-A2.

PD 12-OCT-2000.

PF 06-APR-2000; 2000WO-US009106.

PR 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Caboon RE;

DR WPI; 2000-679375/66.

DR N-PSDB; AAA95276.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.

PS Claim 10; Page 40; 58pp; English.

XX The present sequence is the corn cyclin-dependent kinase inhibitor
 CC (CKI). Its coding sequence was isolated by searching a corn silk cDNA
 CC library for sequences similar to those encoding the CDKI from *Chenopodium*
 CC *rubrum*, *Caenorhabditis elegans* and *Arabidopsis thaliana*. CDKI is involved
 CC in the cell cycle, and may promote or inhibit cell division and growth.
 CC The protein and its coding sequence are useful in the production of
 CC transgenic plants which produce increased or decreased amounts of the
 CC CDKI protein, in the identification of herbicides, in genetic and
 CC physical mapping and in the isolation of the CDKI gene in other organisms

```

XX SQ Sequence 53 AA;
Query Match      8.2%; Score 21; DB 3; Length 53;
Best Local Similarity 100.0%; Pred. No. 5.8e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
DB 27 AKRFASKYNFDFVRGVPLDAG 47

RESULT 6
AAP01940
ID AAP01940 standard; protein; 53 AA.
XX
AC AAP01940;
XX
DT 01-NOV-2001 (first entry)
XX
DE Corn Cyclin dependent kinase inhibitor (CDKI) clone csiln.pk0050.e6.
XX
KW Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; corn;
KW plant growth inhibitor.
XX
OS Zea mays.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR N-PSDB; AAN02390.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 10; Page 40; 58pp; English.
XX
CC The invention describes a novel isolated polynucleotide comprising a
CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
CC of cell division, growth and death. The nucleotide sequences can be used
CC in a vector to transform a host cell to produce the CDKI polypeptide.
CC They can also be used in methods for selecting and obtaining a nucleic
CC acid sequence that encodes CDKI or affects the level of CDKI expression.
CC The encoded protein can be used in a method for evaluating a compound for
CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
CC as herbicides. They can also be used to inhibit plant growth. The
CC polynucleotide sequences can be used in gene mapping and as genetic
CC markers. The sequence is the corn CDKI clone csiln.pk0050.e6 as described
CC in the method of the invention
XX
SQ Sequence 53 AA;
Query Match      8.2%; Score 21; DB 3; Length 53;
Best Local Similarity 100.0%; Pred. No. 5.8e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
DB 27 AKRFASKYNFDFVRGVPLDAG 47

RESULT 7
AAP01940
ID AAP01940 standard; protein; 53 AA.
XX
AC AAP01940;
XX
DT 01-NOV-2001 (first entry)
XX
DE Corn Cyclin dependent kinase inhibitor (CDKI) clone csiln.pk0050.e6.
XX
KW Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; corn;
KW plant growth inhibitor.
XX
OS Zea mays.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR N-PSDB; AAN02390.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 10; Page 40; 58pp; English.
XX
CC The invention describes a novel isolated polynucleotide comprising a
CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
CC of cell division, growth and death. The nucleotide sequences can be used
CC in a vector to transform a host cell to produce the CDKI polypeptide.
CC They can also be used in methods for selecting and obtaining a nucleic
CC acid sequence that encodes CDKI or affects the level of CDKI expression.
CC The encoded protein can be used in a method for evaluating a compound for
CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
CC as herbicides. They can also be used to inhibit plant growth. The
CC polynucleotide sequences can be used in gene mapping and as genetic
CC markers. The sequence is the corn CDKI clone csiln.pk0050.e6 as described
CC in the method of the invention
XX
SQ Sequence 53 AA;
Query Match      8.2%; Score 21; DB 3; Length 53;
Best Local Similarity 100.0%; Pred. No. 5.8e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
DB 27 AKRFASKYNFDFVRGVPLDAG 47

RESULT 8
AAB26249
ID AAB26249 standard; protein; 126 AA.
XX
AC AAB26249;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor #2.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide.

```

```

ABG65674
ID ABG65674 standard; protein; 116 AA.
XX
AC ABG65674;
XX
DT 07-AUG-2003 (revised)
DT 27-AUG-2002 (first entry)
XX
DE Corn ICK 1 protein.
XX
KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX
OS Zea mays.
XX
PN WO200228893-A2.
XX
PD 11-APR-2002.
XX
PF 29-JUN-2001; 2001WO-IB001492.
XX
PR 14-JUL-2000; 2000US-0218471P.
PR 13-OCT-2000; 2000US-0241219P.
XX
PA (CROP-) CROPEDESIGN NV.
XX
PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
DR WPI; 2002-471311/50.
DR N-PSDB; ABK93954.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
PS Claim 48; Fig 4; 141pp; English.
XX
CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents an inhibitor of
CC cyclin dependent kinase (ICK) protein of the invention. (Updated on 07-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 116 AA;
Query Match      8.2%; Score 21; DB 5; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
DB 86 AKRFASKYNFDFVRGVPLDAG 106

RESULT 9
AAB26249
ID AAB26249 standard; protein; 126 AA.
XX
AC AAB26249;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor #2.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide.

```



```

XX OS Zea mays.
XX XX WO200060087-A2.
XX PN 12-OCT-2000.
XX PD
XX PF
XX PF 06-APR-2000; 2000WO-US009106.
XX PF
XX PR 07-APR-1999; 99US-0128192P.
XX XX
XX PA (DUPO ) DU PONT DE NEWMOURS & CO E I.
XX XX
XX PI Klein TW, Weng Z, Cahoon RE;
XX XX
XX DR WPI; 2000-679375/66.
XX DR N-PSDB; AA95280.
XX XX
XX PT Cyclin dependent kinase inhibitor sequences, useful for identifying
XX PT herbicides and plant growth inhibitors.
XX PS
XX PS Claim 10; Fig 1; 58pp; English.
XX CC
XX CC The present sequence is the corn cyclin-dependent kinase inhibitor
XX CC (CDKI). Its coding sequence was isolated by searching a corn silk cDNA
XX CC library for sequences similar to those encoding the CDKI from Chenopodium
XX CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
XX CC in the cell cycle, and may promote or inhibit cell division and growth.
XX CC The protein and its coding sequence are useful in the production of
XX CC transgenic plants which produce increased or decreased amounts of the
XX CC CDKI protein, in the identification of herbicides, in genetic and
XX CC physical mapping and in the isolation of the CDKI gene in other organisms
XX SQ
XX SQ Sequence 126 AA;

Query Match 8.2%; Score 21; DB 3; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
DB 96 AKRFASKYNFDFVRGVPLDAG 116
|||||
|||||

RESULT 9
ABB98759
ID ABB98759 standard; protein; 227 AA.
XX AC ABB98759;
XX DT
XX DT 21-FEB-2003 (first entry)
XX DE
XX DE Maize CKI_D.
XX KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI D; plant; corn;
XX KW crop yield; root size; plant growth; tassel size; ear size;
XX KW male sterility; endoreduplication.
XX XX
XX OS Zea mays.
XX FH Key Location/Qualifiers
XX FT Domain 1..7
XX FT /note= "Conserved domain"
XX FT Misc-difference 111
XX FT /note= "Encoded by CAR"
XX FT
XX PN WO200281623-A2.
XX PN
XX PD 17-OCT-2002.
XX XX
XX PF 06-NOV-2001; 2001WO-US044038.
XX XX
XX PR 07-NOV-2000; 2000US-0246349P.

```

```

XX (PION-) PIONEER HI-BRED INT INC.
XX PA (ARI2-) ARIZONA BOARD OF REGENTS.
XX XX
XX PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX XX
XX DR WPI; 2003-058511/05.
XX DR N-PSDB; ABV74605.
XX XX
XX PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
XX PT involved in cell cycle regulation, and useful for altering cell cycle
XX PT protein content, cell cycle progression, cell number and composition of
XX PT plants.
XX XX
XX PS Claim 12; Page 69; 69pp; English.
XX XX
XX CC The present sequence is maize cyclin-dependent kinase inhibitor (CKI),
XX CC CKI D. The coding sequence for this protein (I) is useful for modulating
XX CC the activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
XX CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
XX CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
XX CC preferably modulating downward is useful for providing differential
XX CC growth in a plant, especially a positive growth advantage and modulating
XX CC CDK activity upward is useful for increasing crop yield, root size, plant
XX CC growth, tassel size and/or ear size. Modulating CDK activity is also
XX CC useful for conferring male sterility and for improving transformation
XX CC frequencies by increasing the number of cells in cell division. CDK
XX CC activity can also be modulated for modulating endoreduplication in the
XX CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
XX CC promoter used is an endosperm-preferred promoter. The cell numbers are
XX CC modulated in one or more tissues of a plant, comprising root, seed,
XX CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
XX CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
XX CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
XX CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
XX CC a second nucleic acid sequence encoding a DNA-binding domain
XX SQ
XX SQ Sequence 227 AA;

Query Match 6.6%; Score 17; DB 6; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 QVGVTRTSRSGAAATGG 39
DB 23 QVGVTRTSRSGAAATGG 39
|||||
|||||

RESULT 10
ABG65670
ID ABG65670 standard; protein; 262 AA.
XX AC
XX AC ABG65670;
XX DT
XX DT 27-AUG-2002 (first entry)
XX DE
XX DE OsiCK 2 protein.
XX KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX OS Oryza sativa.
XX XX
XX PN WO200228893-A2.
XX XX
XX PD 11-APR-2002.
XX XX
XX PF 29-JUN-2001; 2001WO-IB001492.
XX XX
XX PR 14-JUL-2000; 2000US-0218471P.
XX PR 13-OCT-2000; 2000US-0241219P.
XX XX
XX PA (CROP-) CROPDESIGN NV.
XX XX

```


Best Local Similarity 100.0%; Pred. No. 0.63; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAGG 246
Db 205 VRGVPLDAGG 214

RESULT 13

ABG65793
ID ABG65793 standard; peptide; 25 AA.

XX AC ABG65793;

DT 27-AUG-2002 (first entry)

DE Plant potential PEST sequence #9.

XX KW Plant; inhibitor of cyclin dependent kinase; ICK.

OS Oryza sativa.

XX WO200228893-A2.

XX PD 11-APR-2002.

XX PF 29-JUN-2001; 2001WO-IB001492.

XX PR 14-JUL-2000; 2000US-0218471P.

XX PR 13-OCT-2000; 2000US-0241219P.

XX PA (CROP-) CROPEDESIGN NV.

XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatfield Y;

XX DR WPI; 2002-471311/50.

XX PT Novel plant ICK (inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.

XX PS Disclosure; Page 17; 141pp; English.

XX CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents an inhibitor of
CC cyclin dependent kinase (ICK) protein of the invention

XX SQ Sequence 25 AA;

Query Match 3.5%; Score 9; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 SDLESDLAG 189

Db 14 SDLESDLAG 22

RESULT 14

AAP01950

ID AAP01950 standard; protein; 37 AA.

XX AC AAP01950;

XX

DT 01-NOV-2001 (first entry)

XX DE Rice Cyclin dependent kinase inhibitor (CDKI) clone rds2c.pk0088.024.

XX KW Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; rice;
XX plant growth inhibitor.

OS Oryza sativa.

XX Key Location/Qualifiers

FT Misc-difference 33

FT /note= "Encoded by GTCA"

XX PN WO200060087-A2.

XX PD 12-OCT-2000.

XX PF 06-APR-2000; 2000WO-US009106.

XX PR 07-APR-1999; 99US-0128192P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Klein TM, Weng Z, Cahoon RE;

XX DR WPI; 2000-679375/66.

XX DR N-PSDB; AAN02400.

XX PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.

XX PS Claim 10; Fig 1; 58pp; English.

XX CC The invention describes a novel isolated polynucleotide comprising a
CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
CC of cell division, growth and death. The nucleotide sequences can be used
CC in a vector to transform a host cell to produce the CDKI polypeptide.
CC They can also be used in methods for selecting and obtaining a nucleic
CC acid sequence that encodes CDKI or affects the level of CDKI expression.
CC The encoded protein can be used in a method for evaluating a compound for
CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
CC as herbicides. They can also be used to inhibit plant growth. The
CC polynucleotide sequences can be used in gene mapping and as genetic
CC markers. The sequence is the rice CDKI clone rds2c.pk008.024 as described
CC in the method of the invention

XX SQ Sequence 37 AA;

Query Match 3.5%; Score 9; DB 3; Length 37;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAG 245

Db 17 VRGVPLDAG 25

RESULT 15

AAB27252

ID AAB27252 standard; protein; 37 AA.

XX AC AAB27252;

XX DT 17-JAN-2001 (first entry)

XX DE Rice cyclin-dependent kinase inhibitor #3.

XX KW Rice; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
XX cell growth; herbicide.

XX OS Oryza sativa.

XX

FH Key Location/Qualifiers
 FT Misc-difference 34 /note= "encoded by AAGC"
 FT XX
 PN WO200060087-A2.
 XX
 PD 12-OCT-2000.
 XX
 XX 06-APR-2000; 2000WO-US0009106.
 XX
 XX 07-APR-1999; 99US-0128192P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Klein TM, Weng Z, Cahoon RE;
 PI
 XX WPI; 2000-679375/66.
 DR N-PSDB; AAA95286.
 DR
 XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.
 XX
 XX Claim 10; Fig 1; 58pp; English.
 PS
 XX The present sequence is the rice cyclin-dependent kinase inhibitor
 CC (CDKI). Its coding sequence was isolated by searching a rice seed cDNA
 CC library for sequences similar to those encoding the CDKI from Chenopodium
 CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
 CC in the cell cycle, and may promote or inhibit cell division and growth.
 CC The protein and its coding sequence are useful in the production of
 CC transgenic plants which produce increased or decreased amounts of the
 CC CDKI protein, in the identification of herbicides, in genetic and
 CC physical mapping and in the isolation of the CDKI gene in other organisms
 XX
 XX Sequence 37 AA;

Query Match 3.5%; Score 9; DB 3; Length 37;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 237 VRGVPLDAG 245
 Db 17 VRGVPLDAG 25
 |||||

Search completed: October 1, 2004, 17:08:02
 Job time : 128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 17:05:59 ; Search time 18 Seconds
(without alignments)
734.236 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256

Sequence: 1 MGKMKRCRGAAGAEEAAVE.....VRGVPLDAGGRFEWAPVVS1 256

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCUTUS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.1	62	4	US-09-295-996B-4
2	8	3.1	62	4	US-09-295-846B-4
3	8	3.1	62	4	US-09-551-737C-4
4	8	3.1	62	4	US-09-551-738B-4
5	8	3.1	209	4	US-09-685-166A-897
6	8	3.1	254	4	US-09-252-991A-18443
7	8	3.1	358	4	US-09-252-991A-27290
8	8	3.1	492	3	US-09-342-749-2
9	8	3.1	492	4	US-09-691-840-2
10	8	3.1	492	4	US-09-685-166A-895
11	8	3.1	601	1	US-07-676-174A-2
12	8	3.1	601	1	US-08-194-338-2
13	8	3.1	671	4	US-09-252-991A-23641
14	7	2.7	10	4	US-09-755-630B-186
15	7	2.7	15	2	US-08-592-646A-33
16	7	2.7	15	2	US-08-592-646A-35
17	7	2.7	15	2	US-08-592-646A-35
18	7	2.7	15	2	US-08-592-646A-35
19	7	2.7	15	2	US-08-592-646A-36
20	7	2.7	15	2	US-08-592-646A-37
21	7	2.7	15	2	US-08-592-646A-38
22	7	2.7	15	2	US-08-592-646A-39
23	7	2.7	15	4	US-09-165-422-33
24	7	2.7	15	4	US-09-165-422-34
25	7	2.7	15	4	US-09-165-422-35
26	7	2.7	15	4	US-09-165-422-36
27	7	2.7	15	4	US-09-165-422-37
28	7	2.7	15	4	US-09-165-422-38

28 7 2.7 15 4 US-09-165-422-39
29 7 2.7 20 1 US-08-440-861-4
30 7 2.7 33 1 US-08-440-861-54
31 7 2.7 34 1 US-08-433-854-7
32 7 2.7 34 1 US-08-174-745A-7
33 7 2.7 34 2 US-08-195-947-7
34 7 2.7 34 2 US-08-433-885-7
35 7 2.7 34 2 US-08-433-908B-7
36 7 2.7 34 3 US-08-410-614-7
37 7 2.7 37 2 US-08-592-646A-59
38 7 2.7 37 4 US-09-165-422-59
39 7 2.7 48 1 US-07-796-361A-15
40 7 2.7 99 4 US-09-252-991A-32099
41 7 2.7 99 4 US-09-489-039A-9619
42 7 2.7 112 4 US-09-621-976-5911
43 7 2.7 118 4 US-09-732-210-322
44 7 2.7 123 4 US-09-252-991A-30802
45 7 2.7 134 4 US-09-252-991A-22133

ALIGNMENTS

RESULT 1
US-09-295-996B-4
; Sequence 4, Application US/09295996B
; Patent No. 6413530
; GENERAL INFORMATION:
; APPLICANT: BOROVSKY, Dov
; TITLE OF INVENTION: PESTICIDAL PEPTIDES
; FILE REFERENCE: UF-230
; CURRENT APPLICATION NUMBER: US/09/295,996B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-295-996B-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 SAASNSGS 161
DB 38 SAASNSGS 45

RESULT 2
US-09-295-846B-4
; Sequence 4, Application US/09295846B
; Patent No. 6562590
; GENERAL INFORMATION:
; APPLICANT: BOROVSKY, Dov
; TITLE OF INVENTION: Transformed Cells Useful for the Control of Pests
; FILE REFERENCE: UF-223
; CURRENT APPLICATION NUMBER: US/09/295,846B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-295-846B-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 SAASNSGS 161

Mon Oct 4 08:49:12 2004

us-09-993-808b-2.oligo.rai

Db 38 SAASNSGS 45
|||||

RESULT 3

US-09-551-737C-4
; Sequence 4, Application US/09551737C
; Patent No. 6566129
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; APPLICANT: Schlesinger, Yaagov
; APPLICANT: Nauwelaers, Sabine M. I.
; TITLE OF INVENTION: Transformed Cells Useful for the Control of Pests
; FILE REFERENCE: UF-223C1
; CURRENT APPLICATION NUMBER: US/09/551,737C
; CURRENT FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/295,846
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-551-737C-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 154 SAASNSGS 161
|||||

Db 38 SAASNSGS 45

RESULT 4

US-09-551-738B-4
; Sequence 4, Application US/09551738B
; Patent No. 6635265
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: Materials and Methods Useful for the Control of Insect Larvae
; FILE REFERENCE: UF-224C1
; CURRENT APPLICATION NUMBER: US/09/551,738B
; CURRENT FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/296,113
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-551-738B-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 154 SAASNSGS 161
|||||

Db 38 SAASNSGS 45

RESULT 5

US-09-685-166A-897
; Sequence 897, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 897
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-897

Query Match 3.1%; Score 8; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 GAALAAAGL 113
|||||

Db 97 GAALAAAGL 104

RESULT 6

US-09-252-991A-18443
; Sequence 18443, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18443
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18443

Query Match 3.1%; Score 8; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 GAAGQOQA 107
|||||

Db 199 GAAGQOQA 206

RESULT 7

US-09-252-991A-27290
; Sequence 27290, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27290
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27290

Query Match 3.1%; Score 8; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 334 GAALAAAGL 341
|||||

RESULT 8

US-09-342-749-2
; Sequence 2, Application US/09342749
; Patent No. 6166194
; GENERAL INFORMATION:

; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TMRSS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/342,749
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 60/091,044
; EARLIER FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-342-749-2

Query Match 3.1%; Score 8; DB 3; Length 492;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 97 GAALAAAGL 104
|||||

RESULT 9

US-09-691-840-2
; Sequence 2, Application US/09691840
; Patent No. 6444419
; GENERAL INFORMATION:

; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TMRSS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/691,840
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: US/09/342,749
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 60/091,044
; PRIOR FILING DATE: 1998-06-29

; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-691-840-2

Query Match 3.1%; Score 8; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 97 GAALAAAGL 104
|||||

RESULT 10

US-09-685-166A-895
; Sequence 895, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 895
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-895

Query Match 3.1%; Score 8; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 97 GAALAAAGL 104
|||||

RESULT 11

US-07-676-174A-2
; Sequence 2, Application US/07676174A
; Patent No. 5344776
; GENERAL INFORMATION:

; APPLICANT: Vencor, J. Craig
; TITLE OF INVENTION: Octopamine receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARRY & CUSHMAN
; STREET: Eleventh Floor, 1615 L. Street, N.W.
; CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/676,174A
FILING DATE: 19910328
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/79117/KIK
TELEPHONE: (202)861-3000
TELEFAX: (202)822-8944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 601 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-676-174A-2

Query Match 3.1%; Score 8; DB 1; Length 601;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 AGDGGGSC 68
Db 435 AGDGGGSC 442

RESULT 12

US-08-194-338-2
Sequence 2, Application US/08194338
Patent No. 5474898
GENERAL INFORMATION:
APPLICANT: Venter, John C.
APPLICANT: Fraser, Claire M.
APPLICANT: McCombie, William R.
TITLE OF INVENTION: OCTOPAMINE RECEPTOR
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,338
FILING DATE: 08-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 07/676,174
FILING DATE: 28-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH01.001DVI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 601 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-194-338-2

Query Match 3.1%; Score 8; DB 1; Length 601;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 AGDGGGSC 68
Db 435 AGDGGGSC 442

RESULT 13

US-09-252-991A-23641
Sequence 23641, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23641
LENGTH: 671
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23641

Query Match 3.1%; Score 8; DB 4; Length 671;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 ADGAAGQQ 105
Db 522 ADGAAGQQ 529

RESULT 14

US-09-755-630B-186
Sequence 186, Application US/09755630B
Patent No. 6639054
GENERAL INFORMATION:
APPLICANT: ALIBHAI, MURTAZA F.
APPLICANT: ASTWOOD, JAMES D.
APPLICANT: SAMPSON, HUGH A.
APPLICANT: McWHERTER, CHARLES A.
TITLE OF INVENTION: PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS
FILE REFERENCE: 11899.0217.NPUS00 (MOB217)
CURRENT APPLICATION NUMBER: US/09/755,630B
CURRENT FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: US 60/174,669
PRIOR FILING DATE: 2000-01-06
NUMBER OF SEQ ID NOS: 295
SOFTWARE: PatentIn version 3.0
SEQ ID NO 186
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic polypeptide
US-09-755-630B-186

Query Match 2.7%; Score 7; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 FAAREAA 223
|||||
Db 3 FAAREAA 9

RESULT 15
US-08-592-646A-33
; Sequence 33, Application US/08592646A
; Patent No. 5851535
; GENERAL INFORMATION:
; APPLICANT: JOLIVET-REYNAUD, Colette
; TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. Box 19928
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,646A
; FILING DATE: 26-JAN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36923
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 836-6400
; TELEFAX: 703 836-2787
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-592-646A-33

Query Match 2.7%; Score 7; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RGAAGAE 15
|||||
Db 9 RGAAGAE 15

Search completed: October 1, 2004, 17:21:57
Job time : 19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 17:10:40 ; Search time 131 Seconds
(without alignments)
628.859 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEAAVE.....VRGVLDAAGREFAWPVSI 256

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1351062 seqs, 321799191 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	256	100.0	256	9	US-09-993-308-2
2	256	100.0	256	11	US-09-993-808B-2
3	117	45.7	221	12	US-10-425-114-69245
4	117	45.7	263	12	US-10-425-114-63204
5	21	8.2	116	15	US-10-333-006-14
6	17	6.6	227	9	US-09-993-308-6
7	17	6.6	227	11	US-09-993-808B-6
8	17	6.6	262	15	US-10-333-006-10
9	17	6.6	278	12	US-10-425-114-71211
10	17	6.6	417	16	US-10-437-963-195115
11	10	3.9	226	15	US-10-333-006-55
12	10	3.9	354	16	US-10-437-963-122432
13	10	3.9	423	16	US-10-437-963-122443
14	9	3.5	324	12	US-10-107-431-13
15	8	3.1	62	13	US-10-090-696-2
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 69245, A
					Sequence 63204, A
					Sequence 14, Appl
					Sequence 6, Appli
					Sequence 6, Appli
					Sequence 10, Appl
					Sequence 71211, A
					Sequence 195115, A
					Sequence 55, Appl
					Sequence 122432, A
					Sequence 122443, A
					Sequence 13, Appl
					Sequence 2, Appli

16	8	3.1	150	16	US-10-437-963-112674	Sequence 112674, A
17	8	3.1	155	12	US-10-425-114-42206	Sequence 42206, A
18	8	3.1	165	16	US-10-437-963-154569	Sequence 154569, A
19	8	3.1	173	16	US-10-437-963-200561	Sequence 200561, A
20	8	3.1	177	16	US-10-437-963-139291	Sequence 139291, A
21	8	3.1	205	16	US-10-437-963-140830	Sequence 140830, A
22	8	3.1	209	9	US-09-759-143-897	Sequence 897, App
23	8	3.1	209	9	US-09-780-669-897	Sequence 897, App
24	8	3.1	209	9	US-09-822-827-897	Sequence 897, App
25	8	3.1	209	9	US-09-895-793-897	Sequence 897, App
26	8	3.1	209	9	US-09-895-814-897	Sequence 897, App
27	8	3.1	209	13	US-10-012-896-897	Sequence 897, App
28	8	3.1	209	14	US-10-144-678A-897	Sequence 897, App
29	8	3.1	209	14	US-10-294-025-897	Sequence 897, App
30	8	3.1	221	15	US-10-369-493-12041	Sequence 12041, A
31	8	3.1	234	15	US-10-421-138A-128	Sequence 128, App
32	8	3.1	234	15	US-10-374-780A-1242	Sequence 1242, App
33	8	3.1	316	15	US-10-084-848A-116	Sequence 116, App
34	8	3.1	336	12	US-10-282-122A-55536	Sequence 55536, A
35	8	3.1	427	12	US-10-282-122A-60524	Sequence 60524, A
36	8	3.1	491	14	US-10-156-761-13382	Sequence 13382, A
37	8	3.1	492	9	US-09-759-143-895	Sequence 895, App
38	8	3.1	492	9	US-09-759-143-932	Sequence 932, App
39	8	3.1	492	9	US-09-780-669-895	Sequence 895, App
40	8	3.1	492	9	US-09-780-669-932	Sequence 932, App
41	8	3.1	492	9	US-09-879-792-14	Sequence 14, Appl
42	8	3.1	492	9	US-09-822-827-895	Sequence 895, App
43	8	3.1	492	9	US-09-822-827-932	Sequence 932, App
44	8	3.1	492	9	US-09-895-793-895	Sequence 895, App
45	8	3.1	492	9	US-09-895-793-932	Sequence 932, App

ALIGNMENTS

RESULT 1

US-09-993-308-2

; Sequence 2, Application US/099933308

; Patent No. US20020159435A1

; GENERAL INFORMATION:

; APPLICANT: Gordon-Kamm, William J.

; APPLICANT: Lowe, Keith S.

; APPLICANT: Larkins, Brian A.

; APPLICANT: Dilkes, Brian R.

; APPLICANT: Sun, Yuejin

; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof

; FILE REFERENCE: 1146

; CURRENT APPLICATION NUMBER: US/09/993,308

; CURRENT FILING DATE: 2001-11-06

; PRIOR APPLICATION NUMBER: 60/246,349

; PRIOR FILING DATE: 2000-11-07

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 256

; TYPE: PRT

; ORGANISM: Zea mays

US-09-993-308-2

Query Match 100.0%; Score 256; DB 9; Length 256;

Best Local Similarity 100.0%; Pred. No. 3.1e-224;

Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKYMRCRGAAGAEAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAVS 60

DB 1 MGKYMRCRGAAGAEAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAVS 60

QY 61 AGGCGGSCYIHLRSRMFLMAPPQPSVDSVPTPVEAADGAAGCGAALAGLSRCSSTA 120

DB 61 AGGCGGSCYIHLRSRMFLMAPPQPSVDSVPTPVEAADGAAGCGAALAGLSRCSSTA 120

QY 121 SSVNLGIGGGGSHTCRSYDAEAGGHVLVDVSAANSNGSGPDRPRTFTSSRAHGEL 180

Db 121 SSVNLGLGQSGSHTCRSDAAGDGHVLDVSAASNSGSGPDRERETTPSSRAHSEL 180
 QY 181 SDLESDLAGHTKPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
 Db 181 SDLESDLAGHTKPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
 QY 241 PLDAGGRFEWAPVSI 256
 Db 241 PLDAGGRFEWAPVSI 256

RESULT 2
 US-09-993-808B-2
 ; Sequence 2, Application US/09993808B
 ; Publication No. US2004003433A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gordon-Kamm, William
 ; APPLICANT: Lowe, Keith
 ; APPLICANT: Sun, Yuejin
 ; APPLICANT: Dikes, Brian
 ; APPLICANT: Larkins, Brian
 ; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
 ; TITLE OF INVENTION: and Uses Thereof
 ; FILE REFERENCE: 1146
 ; CURRENT APPLICATION NUMBER: US/09/993.808B
 ; CURRENT FILING DATE: 2001-11-06
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 256
 ; TYPE: PRT
 ; ORGANISM: zea mays
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(256)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-993-808B-2

Query Match 100.0%; Score 256; DB 11; Length 256;
 Best Local Similarity 100.0%; Pred. No. 3.1e-224;
 Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGKMRKCRGAGAEVAEVTVGVVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
 Db 1 MGKMRKCRGAGAEVAEVTVGVVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
 QY 61 AGDGGSCYIHLRSMLFMAPPQPPQPSVDSVPTPVEAADGAAGQQAALAGLSCSSTA 120
 Db 61 AGDGGSCYIHLRSMLFMAPPQPPQPSVDSVPTPVEAADGAAGQQAALAGLSCSSTA 120
 QY 121 SSVNLGLGQSGSHTCRSDAAGDGHVLDVSAASNSGSGPDRERETTPSSRAHSEL 180
 Db 121 SSVNLGLGQSGSHTCRSDAAGDGHVLDVSAASNSGSGPDRERETTPSSRAHSEL 180
 QY 181 SDLESDLAGHTKPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
 Db 181 SDLESDLAGHTKPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
 QY 241 PLDAGGRFEWAPVSI 256
 Db 241 PLDAGGRFEWAPVSI 256

RESULT 3
 US-10-425-114-69245
 ; Sequence 69245, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 69245
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI.pep
 US-10-425-114-69245

Query Match 45.7%; Score 117; DB 12; Length 221;
 Best Local Similarity 100.0%; Pred. No. 6e-98;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 140 DAAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHSELSDLESDLAGHTKPSLPAA 199
 Db 105 DAAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHSELSDLESDLAGHTKPSLPAA 164
 QY 200 TPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVSI 256
 Db 165 TPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVSI 221

RESULT 4
 US-10-425-114-63204
 ; Sequence 63204, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 63204
 ; LENGTH: 263
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLM017355F05_FLI.pep
 US-10-425-114-63204

Query Match 45.7%; Score 117; DB 12; Length 263;
 Best Local Similarity 100.0%; Pred. No. 7e-98;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 140 DAAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHSELSDLESDLAGHTKPSLPAA 199
 Db 147 DAAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHSELSDLESDLAGHTKPSLPAA 206
 QY 200 TPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVSI 256
 Db 207 TPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVSI 263

RESULT 5
 US-10-333-006-14
 ; Sequence 14, Application US/10333006
 ; Publication No. US20040019926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
 ; APPLICANT: Peres Bota, Adrian Marius

RESULT 7

Publication No. US20040019926A1
 GENERAL INFORMATION:
 APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.
 APPLICANT: Peres Bota, Adrian Marius
 APPLICANT: Droual, Anne-Marie
 APPLICANT: Mironov, Vladimir
 APPLICANT: Inz, Dirk
 APPLICANT: Hatzfeld, Yves
 TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
 FILE REFERENCE: 1187-13
 CURRENT APPLICATION NUMBER: US/10/333,006
 CURRENT FILING DATE: 2003-01-14
 PRIOR APPLICATION NUMBER: PCT/IB01/01492
 PRIOR FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US 60/218,471
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: US 60/241,219
 PRIOR FILING DATE: 2000-10-13
 NUMBER OF SEQ ID NOS: 59
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 55
 LENGTH: 226
 TYPE: PRT
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: probe or primer
 NAME/KEY: misc feature
 LOCATION: (119)..(119)
 OTHER INFORMATION: Xaa = any amino acid
 US-10-333-006-55

Query Match 3.9%; Score 10; DB 15; Length 226;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAGG 246
 DB 205 VRGVPLDAGG 214

RESULT 12

US-10-437-963-122432
 Sequence 122432, Application US/10437963
 Publication No. US2004012343A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 122432
 LENGTH: 354
 TYPE: PRT
 ORGANISM: Oryza sativa
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_2535C.1.pep
 US-10-437-963-122432

Query Match 3.9%; Score 10; DB 16; Length 354;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAGG 246

US-10-425-114-71211
 Sequence 71211, Application US/10425114
 Publication No. US2004003488A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jingdong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 71211
 LENGTH: 278
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB3078-022-G5_FLI.pep
 US-10-425-114-71211

Query Match 6.6%; Score 17; DB 12; Length 278;
 Best Local Similarity 100.0%; Pred. No. 5.8e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 QVGVTRSRSAATGG 39
 DB 69 QVGVTRSRSAATGG 85

RESULT 10

US-10-437-963-195115
 Sequence 195115, Application US/10437963
 Publication No. US2004012343A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 195115
 LENGTH: 417
 TYPE: PRT
 ORGANISM: Oryza sativa
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1.pep
 US-10-437-963-195115

Query Match 6.6%; Score 17; DB 16; Length 417;
 Best Local Similarity 100.0%; Pred. No. 8.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 ETVQVGVTRSRSA 36
 DB 175 ETVQVGVTRSRSA 191

RESULT 11

US-10-333-006-55
 Sequence 55, Application US/10333006

Db 228 VRGVPLDAGG 237
|||||

RESULT 13

US-10-437-963-122443
; Sequence 122443, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122443
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2536C.1.pep
US-10-437-963-122443

Query Match 3.9%; Score 10; DB 16; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAGG 246
|||||

Db 241 VRGVPLDAGG 250
|||||

RESULT 14

US-10-107-431-13
; Sequence 13, Application US/10107431
; Publication No. US20030224364A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING AND DISTINGUISHING ORTHO
; TITLE OF INVENTION: BIOSYNTHETIC LOCI
; FILE REFERENCE: 3001-7US
; CURRENT APPLICATION NUMBER: US/10/107,431
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Streptomyces mobaraensis
US-10-107-431-13

Query Match 3.5%; Score 9; DB 12; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AAGAEVAAV 19
|||||

Db 27 AAGAEVAAV 35
|||||

RESULT 15

US-10-090-696-2
; Sequence 2, Application US/10090696

This Page Blank (usr)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 09:47:45 ; Search time 2471 Seconds
(without alignments)
3093.775 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MCKYMKRCGAAGAEEAAVE.....VRGVLDPAGRFEPVAVSI 256

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/US9993808/runat_01102004_171702_13035/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9993808 @CNG 1.1 5180 @runat_01102004_171702_13035 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_nam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	820.5	62.9	1197	11	AY108168	AY108168 Zea mays
2	751.5	57.6	659	13	CA074449	CA074449 SCEZAM108
3	744.5	57.1	655	13	CA104638	CA104638 SCJFHR103
4	690.5	53.0	705	13	CA106521	CA106521 SCQGRH101
5	660.5	50.6	661	14	CA256119	CA256119 SCJLFL418
6	656.5	50.3	666	29	CC610974	CC610974 OGVPR67TV
7	655.5	50.3	869	28	BZ743686	BZ743686 OGFAC93TC
8	635.5	48.7	526	14	CD444277	CD444277 EL01N0438
9	590.5	45.3	838	28	CC440801	CC440801 PUHLN53TB
10	582.5	44.7	645	29	CG136784	CG136784 PUICD25TB
11	563.5	43.2	753	14	CB646893	CB646893 OSJNEB09J
12	563	43.2	707	11	AY110319	AY110319 Zea mays
13	522	40.0	585	13	CA104568	CA104568 SCJFHR103
14	518	39.7	642	13	CA104689	CA104689 SCJFHR103
15	517.5	39.7	644	29	CC729680	CC729680 OGBY35TV
16	517	39.6	986	28	CC002416	CC002416 PUGKB53TB
17	514	39.4	584	13	CA108352	CA108352 SCSPHR104
18	512	39.3	510	14	CB929335	CB929335 ABAL 41 H
19	502	38.5	563	13	CA108367	CA108367 SCSPHR104
20	498.5	38.2	987	28	CC002419	CC002419 PUGKB53TD
21	489	37.5	890	29	CG218373	CG218373 OG1DU10TV
22	483	37.0	1010	28	CC390569	CC390569 PUHBT76TB
23	474.5	36.4	531	14	CA297189	CA297189 SCUTSD208
24	458	35.1	697	14	CB646543	CB646543 OSJNEB09B
25	447	34.3	823	29	CC703195	CC703195 OGVBY56TH
26	439.5	33.7	791	29	CC324885	CC324885 OGDND04TV
27	439.5	33.7	917	29	CG265481	CG265481 OG2CM60TV
28	435	33.4	708	28	BZ404098	BZ404098 OGAC123TM
29	435	33.4	795	29	CG339703	CG339703 OG1DH12TH
30	427	32.7	787	28	BZ404090	BZ404090 OGAC123TC
31	399.5	30.6	568	9	AI737717	AI737717 605040C07
32	391.5	30.0	753	29	CC729673	CC729673 OGUSY35TH
33	389.5	29.9	547	28	BH777311	BH777311 fzm5013f0
34	389	29.8	550	12	BG267004	BG267004 1000110B0
35	389	29.8	825	29	CG265472	CG265472 OG2CM60TH
36	375.5	28.8	707	29	CG358355	CG358355 OGWTU21TH
37	371.5	28.5	714	28	BZ642267	BZ642267 OGACU33TM
38	336.5	25.8	456	9	AU172982	AU172982 AU172982
39	307	23.5	467	13	BQ578371	BQ578371 WHE0302 B
40	306.5	23.5	742	14	CB646894	CB646894 OSJNEB09J
41	298.5	22.9	727	14	CB646544	CB646544 OSJNEB09B
42	290.5	22.3	642	28	BZ642255	BZ642255 OGACU33TC
43	286	21.9	837	29	CG030519	CG030519 PUFTA23TB
44	286	21.9	904	29	CG030521	CG030521 PUFTA23TD
45	273.5	21.0	959	29	CC703206	CC703206 OGVBY56TV

ALIGNMENTS

RESULT 1
AY108168
LOCUS Zea mays PC0119368 mRNA sequence.
DEFINITION Zea mays
ACCESSION AY108168
VERSION AY108168.1 GI:21211246
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1197)

AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
AUTHORS 2 (bases 1 to 1197)
Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
1..1197
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Alignment Scores:
Pred. No.: 1.09e-43 Length: 1197
Score: 820.50 Matches: 204
Percent Similarity: 77.01% Conservative: 7
Best Local Similarity: 74.45% Mismatches: 34
Query Match: 62.92% Indels: 29
DB: 11 Gaps: 17

US-09-993-808B-2 (1-256) x AY108168 (1-1197)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 155 ATGGGGAAGTACATCGCAACGACGAGGGGCGCGCGGGGGTGGCGAGTCGAG 214
Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThr---GlyGly 39
Db 215 GTCTCGCAGGTCTGTCGGGTCGGGACGAGGTCCAGTCCGGCGGCGACCGCGCGGT 274
Qy 40 ValAlaLysValAla---ProArgGlyArgAla---ProAla-----52
Db 275 GTCCGGAAGTCTGTCGCGGAGGAGGAGGCGGTCTGCCCGCGCGAAGCGTACG 334
Qy 53 -----GlyGluProAlaAlaValSerAlaGlyLysGlySer---CysTyr 69
Db 335 ACGTGGGGAGCTGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 394
Qy 70 IleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAsp 89
Db 395 ATCCACCTGCGAGCGCATGCTGTCATGGCAGCACCTCAGCAGCAACCGTCG-----448
Qy 90 SerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGly---AlaAla 108
Db 449 GCGGCTCTGACCGCGGTGGAGTCTCT---GGTGGCGACACGAGGGCGGGTGTGGCG 505
Qy 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGly 128
Db 506 CTCGGGGCTGGCTCTCGGCTGTCTCCAGCAGGGGTCTGCTGGTGGAGCTC-----556
Qy 129 GlyGlnArgGlySerHisThrCysArgSerTyrAspAla---AlaGluAlaGlyAsp 147

Db 557 -----GGGGGCCACGCTCGCGCTCGAGCTCGCGCTCGCGAGGTTCGCGGGAT 607
Qy 148 HisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArg 167
Db 608 CAGGTC---CCGATGTCTGTCACCGCGAGCACTCGGGGAGCTCCCGACCGAGAGG 664
Qy 168 ArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeuGluSer 185
Db 665 AGAGAGCAGCAGCGCATGCTCGAGCGGGCGCACGGCGCGAGTCTCAGCATCTTGAGTCG 724
Qy 186 AspleuAlaGly---HisLysThrGly---ProSerLeuProAlaAlaThrProAla---202
Db 725 GATCTGTGTGGGGCGGAGAGACTGGCTGCTGTCGCGCGGCGACACACATCGGTT 784
Qy 203 AlaGluLeuLeuValProProAlaHisGluIleGlnGluPheAlaAlaAlaGluAla 222
Db 785 GCGGAGCTGATGTCGCGCGCAGCAGGAGATCAGGAATTCCTCGCGCGCGCGAGGCG 844
Qy 223 AlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeu 242
Db 845 GCCCATGCCAAAGCTTTGCTTCCAGTACACTTCGACTTCGCTCGCGCGGTGCCCTC 904
Qy 243 AspAlaGlyGlyArgPheGluTtpAlaProValValSerIle 256
Db 905 GACGCC---GCCCGGTTTCGAGTGGACGCCAGGGGTTCAGCATC 943

RESULT 2
CA074449 659 bp mRNA linear EST 23-SEP-2003
LOCUS SCEZAM1081D06.g AM1 Saccharum officinarum cDNA clone SCEZAM1081D06
DEFINITION 5', mRNA sequence.
ACCESSION CA074449
VERSION CA074449.1 GI:34926722
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum

REFERENCE 1 (bases 1 to 659)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 081 row: D column: 06
seq primer: T7 Promoter Primer.

FEATURES
source
1..659
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEZAM1081D06"
/lab_host="DH10B"
/clone_lib="AM1"
/note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be

obtained at <http://sucet.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:

Pred. No.: 1.61e-39 Length: 659
Score: 751.50 Matches: 173
Percent Similarity: 84.21% Conservative: 3
Best Local Similarity: 82.78% Mismatches: 14
Query Match: 57.63% Indels: 19
DB: 13 Gaps: 11

US-09-993-808B-2 (1-256) x CA074449 (1-659)

QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaAlaVal 19
Db 59 ATGGGGAAGTACATCGCAGCAAGTCAGGGCGCGGGCGAGAGTTCGCCCGCGTC 118
QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThr---Gly 38
Db 119 GAGTCAACGAGGTCTGTCGGCGTCCGACAGGTCCAGGTCCGGCGCGCGCGGGGT 178
QY 39 GlyValAlaLysValAlaProArgArgGlyArgAla-----ProAla 52
Db 179 GGTGTGGCAAGTCTGCACCGAGGAGGAAGCGCTGGCGCGCGCAACGTGGCGCG 238
QY 53 GlyGluProAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 239 GGGGAGCT---GCTGCGTGGCGTGGAGGGACGGCGGAGCTGCTACATCCATCTG 295
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 296 CGCAGCGCATGCTGTTTCATGGCACCCTCAGCGCGCGCG-----TCTGTTCCG 346
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
Db 347 ACCCGCGGAGGTGCTGTGTGA-----GGACAGCAGCGCGCGCGCTCGCGCGCG 400
QY 113 LeuSerArgCysSerThrAlaSerSerValAsnLeuGlyLeuGlyGlyArgGly 132
Db 401 CTCTCGGTGCTCCAGCAGCGGTCTGCTGGTGGAGCTCGGG---GTCATCAGCAGG 457
QY 133 SerHisThrCysArgSerThrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
Db 458 AGCCAGCGCTCGCGCTCC---GACGCTGCGGAGGTGGCGGNGATCAGCGC---CGGAT 511
QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
Db 512 GTCTCG---GCGAGCACTCGGGGAGCGGCCCGAGCGGAGGAGAGAGAGAGAGAGCA 568
QY 173 SerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys--- 191
Db 569 TCGAGCGGGCGCAGCGGAGCTCAGCGATCTGGAGTCTGGAGTCTGGCGGGGCAAGAAT 628
QY 192 ThrGlyProSerLeuProAlaAlaThr 200
Db 629 ACTGGCGGTCTGTCGGCGGCAACG 655

RESULT 3

CA104638/c

LOCUS

DEFINITION

SCUFRH1034B05.b HRI Saccharum officinarum cDNA clone SCUFRH1034B05

3', mRNA sequence.

CA104638

CA104638.1 GI:34957945

EST.

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Saccharum.

1 (bases 1 to 655)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

JOURNAL
COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

<http://www.bccccenter.fcav.unesp.br>

Plate: 034 row: B column: 05

Seq primer: SP6 Promoter primer.

Location/Qualifiers

FEATURES

source

1..655

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCUFRH1034B05"

/lab_host="DH10B"

/clone_lib="HRL"

/note="Organ: seedlings inoculated with *Herbaspirillum*

rubrisubalbicans; Vector: pSport1; Site 1: SalI; Site 2:

NotI; An unidirectional cDNA library generated from

[seedlings inoculated with *Herbaspirillum*

rubrisubalbicans]. cDNA was prepared from polyA+ mRNA

using SuperScript Plasmid System Kit (Invitrogen). The

double-strand cDNAs were fractionated in a sepharose

CL-2B 40cm-columns and fragments sizing between 0.8 and

1.5 Kb were directionally cloned into the vector. Details

of each source of RNA and library construction can be

obtained at <http://sucet.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:

Pred. No.: 4.57e-39 Length: 655
Score: 744.50 Matches: 172
Percent Similarity: 80.09% Conservative: 9
Best Local Similarity: 76.11% Mismatches: 26
Query Match: 57.09% Indels: 19
DB: 13 Gaps: 11

US-09-993-808B-2 (1-256) x CA104638 (1-655)

QY 1 MetGlyLysTyrMet---ArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaVal 19
Db 648 ATGGGAAGTAACTCCAGTCCCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 589
QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db 588 GAGCTCAGCAGATCTGCGCGTCCGAGCAGATCCAGATCCGCGCGCGCGCGCGGGGT 529
QY 40 ---ValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 528 GCTGTGGCAAGATCGCACCGAGGAGGAAGAGCGCTGGCGCGCTCCACAGTGGCGGCG 469
QY 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 468 GGGGAGCT---GCTGCGCTGGCGCTGGAGGGGACGGCGGCGGCTCGCGCGCGGG 412
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 411 CGCAGCGCATGCTGTTTCATGGCACCCTCAGCGCGCGCTGCTACATCCATCTG 361
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
Db 360 ACCCGCGGAGGTGCT-----GGTGCAGACACAGCGCGCGCGCTCGCGCGCGGG 307
QY 113 LeuSerArgCysSerThrAlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGly 132
Db 306 CTCTCGGTGCTCCAGCAGCGGTCTGCTGGTGGAGCTCGGG---GGTATCAGCGCAGG 250
QY 133 SerHisThrCysArgSerThrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152

Db 249 AGCCACGCTGCTCC---GAGCTGGAGGTTGGCGGGATCAGCC---CCGAT 196
 Qy 153 ValSerAlaSerAenSerGlySerGlyProAspArgGluArgGluThrPro 172
 Db 195 GTCTCG---GGAGCAACTCGGGAGCGGCCACACCGGAGAGAGACACGCCA 139
 Qy 173 SerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys--- 191
 Db 138 TCAGACCGGGCGGCGGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCAACAAGACT 79
 Qy 192 ThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuValProAlaHis 211
 Db 78 ACTCCCGGCTGCTGGCGGGCGGCAACGTCGCTGGCAGCGGATCGTGCCACGACCAA 19
 Qy 212 GluLeuGlnGluPhe 217
 Db 18 GAGATCGAGGAGTCTTC 1

RESULT 4
 CA106521
 LOCUS SCQHR1013C07.g HRI Saccharum officinarum cDNA clone SCQHR1013C07
 DEFINITION 5', mRNA sequence.

ACCESSION CA106521
 VERSION CA106521.1 GI:34959828
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 JOURNAL clade; Panicoideae; Andropogoneae; Saccharum.
 COMMENT 1 (bases 1 to 705)

REFERENCE
 AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 TITLE The libraries that made SUCEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089

Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccccenter.fcav.unesp.br
 Plate: 013 row: C column: 07
 Seq primer: T7 Promoter Primer.

FEATURES
 Location/Qualifiers
 1..705
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCQHR1013C07"
 /lab_host="DH10B"
 /clone_lib="HRI"
 /note="Organ: seedlings inoculated with Herbaspirillum
 rubrisubalbicans; Vector: pSport1; Site_1: SalI; Site_2:
 NotI; An unidirectional cDNA library generated from
 [seedlings inoculated with Herbaspirillum
 rubrisubalbicans]. cDNA was prepared from polyA+ mRNA
 using SuperScript Plasmid System Kit (Invitrogen). The
 double-strand cDNAs were fractionated in a sepharose
 Clu-2B 40cm-columns and fragments sizing between 0.8 and
 1.5 Kb were directionally cloned into the vector. Details
 of each source of RNA and library construction can be
 obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.67e-35 Length: 705
 Score: 690.50 Matches: 162
 Percent Similarity: 81.95% Conservative: 6

Best Local Similarity: 79.02% Mismatches: 18
 Query Match: 52.95% Indels: 19
 DB: 13 Gaps: 11

US-09-993-808B-2 (1-256) x CA106521 (1-705)

Qy 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaGlyAlaGluValAlaVal 19
 Db 113 ATGGGGAAGTACATGCGCAGCACTGCAGGGGGCGCGGGGAGGAGTCCCGCGCTC 172
 Qy 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThr---Gly 38
 Db 173 GAGTCTCAGCAGGTCGTCGGCGCTCCGACGAGGTCCAGGTCCGCGGGCGGACCGGGGT 232
 Qy 39 GlyValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
 Db 233 GGTGTGGCGAAGGTCCGACCGAGGAGGAAGGGCTGGCGCGGCCCAACGTGGCGCG 292
 Qy 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
 Db 293 GGGGAGGCT---GCTGGCGTGGCGCTGGAGGGGACGGCGGAAGCTGTACATCATCTG 349
 Qy 73 ArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValPro 92
 Db 350 CGCAGCGCATGCTGTTCATGCGACCGCTCAGCGCGAGCGG-----TCGTTCCG 400
 Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
 Db 401 ACGCGGGGAGGCTGCT-----GGTCAGGACANACAGGGCGCGGCTCGCGCGGG 454
 Qy 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGly 132
 Db 455 CTCTCGGCTGCTCCAGCAGCGGCTGCTGGTGGAGCTCGGG---GGTCATCAGCGCAAG 511
 Qy 133 SerHisThrCysArgSerTyrAspAlaGluAlaGlyGlyAspHisValLeuValAsp 152
 Db 512 AGCCACGCTCGCGCTCC---GAGCTCGGAGGTGGCGGGATCAGCC---CCGAT 565
 Qy 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrPro 172
 Db 566 GTCTCN---GGAGCAACTCGGGAGCGCGCCCAACCGGAGAGAGACGACGCC 622
 Qy 173 SerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys--- 191
 Db 623 TCAGACCGGGCGCACGGCAGCTCAACGATCTGGAGTTCGATCCGCGGGGACAAAGATT 682
 Qy 192 ThrGlyProSerLeu 196
 Db 683 ACGCGCGGTCGTTG 697

RESULT 5

CA256119

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA256119 661 bp mRNA linear EST 26-SEP-2003
 SCULFL4187B06.g FL4 Saccharum officinarum cDNA clone SCULFL4187B06
 5', mRNA sequence.

CA256119

CA256119.1 GI:35941705

EST

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum.

1 (bases 1 to 661)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bccc.net.br>
Plate: 187 row: B column: 06
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .661
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCULF4187B06"
/lab_host="DH108"
/clone_lib="FL4"
/note="Organ: Developed inflorescence and rachis (20cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [developed] inflorescence and rachis (20cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucet.lad.ic.unicamp.br/public>

FEATURES

source

QY 173 SerSerArgAlaHisGlyGluLeuSerAsp 182
Db 631 TTTAACCCGGCGACGCGGAGCTTAACGAT 660
RESULT 6
CC610974/c
LOCUS OGVFR67TV.ZM.0.7.1.5.KB.Zea.mays.genomic.clone.ZMMBMA0551L14,
genomic survey sequence.
ACCESSION CC610974
VERSION CC610974.1 GI:31972395
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 666)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
JOURNAL Other GSSs: OGVFR67TH
COMMENT Contact: Cathy Whitelaw
TIGR
712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1. .666
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0551L14"
/clone_lib="ZM.0.7.1.5.KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES

source

US-09-993-808B-2 (1-256) x CA256119 (1-661)
QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaVal 19
Db 121 ATGGGAAGTACATCGCAGCAAGTGCAGGGGGCGCGCGGAGGAGTGCAGCGGCTC 180
QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThr---Gly 38
Db 181 GAGGTACCGCAGGTCGTGCGCGTCCGACAGGTCCAGGTCCGCGCGGCGGCGGT 240
QY 39 GlyValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 241 GGTGTGCGGAGTGTCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 301 GGGGAGCCT---GCTGCGGTGCGCGTGGNAGGAGCGCGGAGGAGTGTACATCCATCTG 357
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 358 CCGAGCGCATGTGCTGTATGCGACCGCGCTCAGCGCGGAGCGG-----TCGCTCCG 408
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlnGlnGlnGlnGlnGln 112
Db 409 ACCCGCGGAGGAGTGTCT-----GTTGAGGAGACAGCAGCGCGCGGCTCGCGCGG 462
QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlyGlyGly 132
Db 463 CTCTCGGTTGCTCCAGCAGGCGTGTGCTGCTGAGAGTCTGCGG---GGTATCAGCGCAGG 519
QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyAspHisValLeuValAsp 152
Db 520 AGCCAGCGCTCGCGCTCC---GACGCTGCGAGAGTGTGGCGGGATCAGCGC---CCGGAT 573
QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
Db 574 GTCTCG---GCGAGCAACTCGGGGAGCGGGCCAGACCGGAGAGAGAGAGAGAGAGAG 630

ORIGIN

Alignment Scores:
Pred. No.: 2,76e-33 Length: 666
Score: 656.00 Matches: 142
Percent Similarity: 81.14% Conservative: 0
Best Local Similarity: 81.14% Mismatches: 7
Query Match: 50.31% Indels: 26
DB: 29 Gaps: 5
US-09-993-808B-2 (1-256) x CC610974 (1-666)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 467 ATGGGAAGTACATCGCAGCAAGTGCAGGGGGCGCGGAGGAGTGCAGCGCGGCTG 408
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 407 GTTACGAGGTGTGTCGCGTCCGACAGGTCCAGGTCCGCGCGGCGGAGCGCGGCTG 348
QY 41 AlaLys---ValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
Db 347 GCGAAGTGTGTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 288
QY 60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
Db 287 GCGCGTGTGGGAGCGCGGAGGAGTGTACATCCACCTCGAAGCGCGCATGCTGTTTCATG 228
QY 80 AlaProGlnProGlnProSerValAspSerValProThrPro---ValGluAlaAla 98
Db 227 GCACCGCTCAGCGCGGAGCGG-----TCGGTTCGACCCCGCGGAGGAGGCTGCT 177

```

QY 99 AspGlyAlaLagGlyGlnGlnGlyAlaAlaLeuAlaLagGlyLeuSerArgCysSerSer 118
Db 176 GATGGCGCTGCAGACAGCAGCGGGCGGTCTCGCGCGGGCTCTCGCGTGTCTCCAGC 117
QY 119 ThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSer 138
Db 116 ACGCGCTCGTGGTGAACCTTGGGGTTGGGGTTCAGCGCGGAGCCACACCTGCGCTCC 57
QY 139 TyrAspAlaAlaGluAlaGlyCysAspHisValLeuValAspValSerAlaAlaSerAsn 158
Db 56 GACGACGCTGCAGAGGTG----- 39
QY 159 SerGlySerGlyProAspArgGluArg---ArgGluThrThrPro 172
Db 38 -----AGAGAGAGAGAGAGAGAGCCAAACCT 12

RESULT 7
BZ743686/c 869 bp DNA linear GSS 03-MAR-2003
LOCUS OGFAC93TC ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0228P17,
DEFINITION genomic survey sequence.
ACCESSION BZ743686
VERSION BZ743686.1 GI:28723396
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 869)
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGFAC93TM
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
    source
        1..869
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBma0228P17"
            /clone_lib="ZM 0.7 1.5 KB"
            /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 4,18e-33 Length: 869
Score: 655.50 Matches: 136
Percent Similarity: 93.79% Conservative: 0
Best Local Similarity: 93.79% Mismatches: 4
Query Match: 50.27% Indels: 5
DB: 28 Gaps: 3

US-09-993-808B-2 (1-256) x BZ743686 (1-869)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 446 ATGGGGAGTACATGCGCAGTGCAGGGGGCGCAGGCGCGAGGTGCGCGCGCTCGAG 387
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40

```

```

Db 386 GTTACGAGGTCGTGCGGTCGCCACAGAGTCCAGGTCCGCGCGGGACCGCGGTGTC 327
QY 41 Alalys---ValAlaProArgArgLeuArgAlaProAlaGlyGluProAlaAlaVal 59
Db 326 GCGAAGTGTGTCGCCCGGAGGAGAGAGGGCGCGCGGGGGAGCTGTGTCGCCGTG 267
QY 60 SerAlaGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
Db 266 GCGCTGTGTGGGACGCGGAAAGTGTATCATCTCACCTGCGAAGCGCATCTGTTCATG 207
QY 80 AlaproProGlnProGlnProSerValAspSerValProThrPro---ValgluAlaAla 98
Db 206 GCACGCGCTCAGCGCAGCGC-----TCGGTTCCGACCCCGCGCGGAGGCTGCT 156
QY 99 AspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSer 118
Db 155 GATGGCGCTGCAGACAGCAGCGCGGTCTCGCGCGGGCTCTCGCGTGTCTCCAGC 96
QY 119 ThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSer 138
Db 95 ACGCGCTGTCGTGTAACCTTGGGGTTGGGGTTCAGCGCGGAGCCACCTGCGCTCC 36
QY 139 TyrAspAlaAlaGlu 143
Db 35 GACGACGCTGCAGAG 21

RESULT 8
CD444277 526 bp mRNA linear EST 03-JUN-2003
LOCUS EL01N0438A02.b Endosperm_4 Zea mays cDNA, mRNA sequence.
DEFINITION CD444277
ACCESSION CD444277 GI:31359920
VERSION CD444277.1
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 526)
Lai,J., Dev,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
Messing,J.
Sequencing of the maize endosperm ESTs
Unpublished (2002)
Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
Location/Qualifiers
    1..526
        /organism="Zea mays"
        /mol_type="mRNA"
        /cultiivar="W22"
        /db_xref="taxon:4577"
        /tissue_type="Endosperm of 7-23DAP"
        /clone_lib="Endosperm 4"
        /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
        XhoI"

FEATURES
    source
        1..526
            /organism="Zea mays"
            /mol_type="mRNA"
            /cultiivar="W22"
            /db_xref="taxon:4577"
            /tissue_type="Endosperm of 7-23DAP"
            /clone_lib="Endosperm 4"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 4,78e-32 Length: 526
Score: 635.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.70% Indels: 0
DB: 14 Gaps: 0

US-09-993-808B-2 (1-256) x CD444277 (1-526)

```

QY	1	MetGlyIysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAaValGlu	20
Db	146	ATGGGGGAAGTACATGCGCAAGTCAGGGGGCGCGCAGCGCGGAGGTCGCCGCCGTCGAG	205
QY	21	ValThrGlnValValGlyValAlaThrArgSerAlaAlaAaThrGlyClyVal	40
Db	206	GTTACGACGAGTCGTGCGCGTCCGGACGAGGTCACAGTCCGCGCGCGCACCGCGCGGTGTC	265
QY	41	AlaLysValAlaProArgLysArgAlaProAlaGlyGluProAlaAaValSer	60
Db	266	GCGAAGGTCGCCCGCGAGGAGGAGGCGCGCGGGGGAGCCTGCTGCCCGCTGAGC	325
QY	61	AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla	80
Db	326	GCTGTGTGGGACCGCGGAAGCTGCTACATCCACCTGCGTAGCCGCATGCTGTTCATGSCA	385
QY	81	ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAaAspGly	100
Db	386	CCGGCTCAGCCCGACCGCGTCGGTGTGACTTCGGTTCGACCCCGGTGGAGGCTGCTGATGCGC	445
QY	101	AlaAlaGlyGlnGlnGlyValAaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla	120
Db	446	GCTGCAGACAGCAGGGCGCGCGCTCGCTCGCGGGCTCTCGCGTGTGCTCCAGACACGCGG	505
QY	121	SerSerValAsnLeuGlyLeu	127
Db	506	TCGTCCTGTGAACTTGGGCTTG	526

RESULT 9				
CC440801/c				
LOCUS	CC440801	838 bp	DNA linear	GSS 20-MAY-2003
DEFINITION	PURLN53TB_ZM_0.6_1.0 KB Zea mays genomic clone ZMMBTA478J10,			
	genomic survey sequence.			

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 838)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
 Bennetzen, J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other GSIDs: PURLN53TD
 Contact: Cathy Whitelaw
 TIGR

ORIGIN

Alignment Scores:

Pred. No.:	6.93e-29	Length:	838
Score:	530.50	Matches:	123
Percent Similarity:	94.03%	Conservative:	3
Best Local Similarity:	91.79%	Mismatches:	3

```

Query Match: 45.28% Indels: 5
DB: 28 Gaps: 3
US-09-993-808B-2 (1-256) x CC440801 (1-838)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
   |||||
Db 390 ATGGGAAGTACATGCGCAAGTGCAGGGGGCCGAGCGCGGAGGTCCGCGCGCTCGAG 331
   |||||

Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGlyVal 40
   |||||
Db 330 GTTACGCAGTCTGTCGGGCTCCGCACAGGTTCCAGTCCGCGCGCGACCGCGCGTGC 271
   |||||

Qy 41 AlaLys---ValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaVal 59
   |||||
Db 270 GCGAAGGTCTGTCGCCGCCGAGGAGGAGGGCGCGCGGGGGAGCTGCTGCCTCCGTG 211
   |||||

Qy 60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
   |||||
Db 210 GGCCTCTGTGGGGACGGCGGAAGCTGTACATCCACCTGCTAGCCGATGCTGTTCATG 151
   |||||

Qy 80 AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
   |||||
Db 150 GCACCGCTCAGCCGCGACCCG-----CCGTCTGGTTCCGACCCCGCGAGGCTGCTCAT 97
   |||||

Qy 100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThr 119
   |||||
Db 96 GGCCTCCAGACACAGCGGCGCGCTCGCGCGCGGCTCTCGCGTTGCTCCAGCAGC 37
   |||||

Qy 120 AlaSerSerValLeuLeuGlyLeuGlyGlyGlnArgGlySer 133
   |||||
Db 36 GCGTCTGTCGTGTGCAC-----GTGGGGGGTTCAGCGCGGGAGC 1
   |||||

```

RESULT 10	ACCESSION	ORGANISM
CG136784	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

REFERENCE

1 (bases 1 to 645)

Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

ORIGIN

Alignment Scores:

Pred. No.: 1.65e-28 Length: 645
 Score: 582.50 Matches: 130
 Percent Similarity: 60.55% Conservative: 2
 Best Local Similarity: 59.63% Mismatches: 3
 Query Match: 44.67% Indels: 83
 DB: 29 Gaps: 4

US-09-993-808B-2 (1-256) x CG136784 (1-645)

Qy 75 ArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValProThrPro 94
 Db 2 CGCATGCTGTTCATGGCAGCCCTCAGCGGAGCG-----CGTCGGTTCAGCCCG 55
 Qy 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 114
 Db 56 GCGGAGGCTGTAAGGCTGTCAGGACAGCGGCGGCGGCTGCGCGGCTCTCG 115
 Qy 115 ArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGlySerHis 134
 Db 116 CGTTGCTCCAGCAGCGGCTGCTGCTGCAC-----GTGGGGGGTTCAGCGGAGCCAC 169
 Qy 135 ThrCysArgSerTyrAspAlaAlaGlu-----143
 Db 170 ACCTGCCGCTCCGACGAGCTGCAGAGGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 229
 Qy 143 -----143
 Db 230 GAATTATTCGAATTCATAAAACGACCGAGATCTCCAGCGCTACTCGTTGCTGAATTCGT 289
 Qy 144 -----AlaGlyGlyAspHisValLeu 150
 Db 290 GGCTTTCTTTTGGTTTGGTCTCTCTACGACGTAACAGGCTGCGGGGATCAGTCTCG 349
 Qy 151 ValAspValSerAlaAlaSerAsnSerClySerGlyProAspArgGlu-----166
 Db 350 GTGGATGCTCGCGGGGAGAGAACTCCGGGAGCGGCCAGACCCGAGAGGTTGGAGAA 409
 Qy 166 -----166
 Db 410 TCTCCCTTGTCTCGCGATCTCTGCTGAATTCATGCTTCGCGGCTCTCTGACGTCG 469
 Qy 167 -----ArgArgGluThrThrProSerSerArg 175
 Db 470 TTCGCGCGCGCGCTGCTGGTTTGGAAATTCAGCGGAGAGAGACGCCATCAGCGCG 529
 Qy 176 AlaHisGlyGluLeuSerAspLeuGlySerAspLeuAlaGlyHisLysThrGlyProSer 195
 Db 530 GCGACGCGGAGCTCAGCGATCTGAGTCGATCTGGCGGGGCAAGACTGGCCCGTCG 589
 Qy 196 LeuProAlaAlaThrProAlaAlaGluLeuValProAlaHisGluLeu 213
 Db 590 CTACCGCGGCAACCGCGGCTCGGAGCTGATCTGTCGCGCCAGCACAGAGATC 643

RESULT 11

CB646893

LOCUS

DEFINITION OSJNEB09J11.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA

EST. OSJNEB09J11 5', mRNA sequence.

ACCESSION

CB646893

VERSION

CB646893.1

KEYWORDS

EST.

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 753)

Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,

Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.

Large-scale identification of ESTs involved in the interaction

between rice and Magnaporthe grisea

unpublished (2003)

COMMENT

Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu

FEATURES

source
 1..753
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEB09J11"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEB"
 /note="vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
 XhoI; 24 hrs after inoculation with Rice Blast (Che
 80661)"

ORIGIN

Alignment Scores:
 Pred. No.: 3.48e-27 Length: 753
 Score: 563.50 Matches: 140
 Percent Similarity: 65.43% Conservative: 19
 Best Local Similarity: 57.61% Mismatches: 65
 Query Match: 43.21% Indels: 19
 DB: 14 Gaps: 9

US-09-993-808B-2 (1-256) x CB646893 (1-753)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
 Db 46 ATGGCAAGTACATGAGGAAGTTCAGGGGGGCCAGGGGAGAGATTGGCGCCATGAG 105
 Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly--- 39
 Db 106 GTCACGAGGTGTGGCTCGGACGAGGTCTGAGTCTGCGGCGGCGGCGGCGGCGGCG 165
 Qy 40 -----ValAlaLysValAlaProArgLysArgAlaProAlaGlyGlu 54
 Db 166 ACGACGAGGTGAAGGCGGCTCGCGGCGCTCCACGAGGAGGAGAGGCGCTGCTGCGG 225
 Qy 55 ProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSer 74
 Db 226 ACGCGGTCTGGGGACTACTCGCGTGACGCGGGAGCTGCTACCTCCAGCTGAGGAGC 285
 Qy 75 ArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValProThrPro 94
 Db 286 CGCATGCTGTTCATGGCCCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 342
 Qy 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 114
 Db 343 GCGGAGGCGCGG---GGTTCGGGAACGAGGCGGCGGCGGCGGCGGCGGCTGCGCTCTCG 399
 Qy 115 ArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlyGlyGlnArgGlySerHis 134
 Db 400 CGTTCTCCAGCAGCGGCTGCTCGCTGGAC-----GCGGCGGCTCAGGACAGGAGCCTC 453
 Qy 135 ThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
 Db 454 GCGTCCGCTCC---GAGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 504
 Qy 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrProSerSer 174
 Db -----

PCR Primers
 FORWARD: gta aac cga cgg cca gta
 BACKWARD: gga aac agc tat gac cat g
 Plate: 09 row: J column: 11
 Seq primer: gta aac cga cgg cca gta

Location/Qualifiers


```

Db      505  TCCGCGAGCGACTCGCGGAGCGCGCTGAGCCGCGAGAGAGAGAACTCCATCAAGC 564
QY      175  ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAla---GlyHisLysThrGly 193
Db      565  TTTCTCCCGGAGGTGAGCATCTGGAGTGGATCTGGCTGGAGGACAGAGCGCAGC 624
QY      194  ProSerLeuPro---AlaAlaThrProAlaAlaGluLeuIle-----ValPro 208
Db      625  CGTCACACTACTCTGCGGCAACAGCTCAGCACAGCAAGCCAGCGCGCGAAGATTCCG 684
QY      209  ProAlaHisGluLeuGlnGluPheAlaAlaAlaGluAlaGlnAlaLysArgPhe 228
Db      685  CGGCGCGCGAGATCGAGGCTTCTTCGCGCGCGCGGCGGAGAGAGCTGAGGCCAAGCGCTTC 744
QY      229  AlaSerLys 231
Db      745  GCGGCCAAG 753

RESULT 12
LOCUS   AY110319
DEFINITION Zea mays CL528_1 mRNA sequence.
ACCESSION AY110319
VERSION   AY110319.1 GI:21214635
KEYWORDS  HTC.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 707)
AUTHORS   Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S.,
          Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
          Maize Mapping Project/DuPont Consensus Sequences for Design of
          Zea mays
          Overgo Probes
          Unpublished (2002)
          2 (bases 1 to 707)
          Coe, E.H.
          Direct Submission
          Submitted (25-APR-2002) Maize Mapping Project, University of
          Missouri, Columbia, MO 65211, USA
          If you are interested in getting corresponding physical clones,
          these are publicly available from ZmDB and may be found by BLAST
          searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
          www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
          maize cDNA sequences is either Virginia Walbot, Stanford or Pat
          Schnable, Iowa State, then clones may be requested from ZmDB:
          www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source           1..707
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:632312"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Consensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"

ORIGIN
Alignm Scores:      3.46e-27      Length:      707
Pred. No.:          563.00      Matches:     110
Score:              88.80%      Conservative: 1
Percent Similarity: 88.00%      Mismatches:  14
Best Local Similarity: 43.17%   Indels:      0
Query Match:        11         Gaps:        0
DB:
US-09-993-808B-2 (1-256) x AY110319 (1-707)

```

```

QY      132  GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyAspHisValLeuVal 151
Db      10  GGAGGACACACTGCCGCTCCAGCAGCTGCAGAGCTGGCGGGGATCAGCTCTGGTG 69
QY      152  AspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThr 171
Db      70  GTGGATGTCTCGCGAGCAACTCCGCGAGCGGCCAGACCGCGAGAGGAGAGACGACG 129
QY      172  ProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys 191
Db      130  CCATCGAGCGCGCGCACGCGAGCTCAGCGATCTGGAGTCGATCTGGCGGGGCAAG 189
QY      192  ThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIleValProAlaHis 211
Db      190  AITGGCCGCTCGCTACCGCGGCACCGCGCTGCGGAGCTGATCTGCCGCGCAGCAC 249
QY      212  GluIleGlnGluPheAlaAlaGluAlaGlnAlaLysArgPheAlaSerLys 231
Db      250  GAGATCCAGAGTTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 309
QY      232  TyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluThrAla 251
Db      310  TACAACCTTCGACTTCGTCGCGCGCGTGCCTCGACGCGCGCGCGCTTCGAGTGGCG 369
QY      252  ProValValSerIle 256
Db      370  CCGGTGGTCAGCATC 384

CA104568      585 bp      mRNA      linear      EST 23-SEP-2003
SCJFHR1034B05.g HRI Saccharum officinarum cDNA clone SCJFHR1034B05
5', mRNA sequence.
CA104568
CA104568.1 GI:34957875
EST.
SOURCE        Saccharum officinarum
ORGANISM      Saccharum officinarum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE 1 (bases 1 to 585)
AUTHORS       Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
          The libraries that made SUCEST
          Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
          Contact: Arruda P
          Centro de Biologia Molecular e Engenhariaia Genetica
          Universidade Estadual de Campinas
          Caixa Postal 6010, 13083-970, Campinas SP, Brazil
          Tel: 55 19 3788 1137
          Fax: 55 19 3788 1089
          Email: parruda@unicamp.br
          Clone distribution: clone distribution information can be found
          through the Brazilian Clone Collection Center (BCCC) at
          http://www.bcccenter.fcav.unesp.br
          Plate: 034 row: B column: 05
          Seq primer: T7 Promoter Primer.
          Location/Qualifiers
             1..585
             /organism="Saccharum officinarum"
             /mol_type="mRNA"
             /db_xref="taxon:4547"
             /clone="SCJFHR1034B05"
             /lab_host="DH10B"
             /clone_lib="HRI"
             /note="Organ: seedlings inoculated with Herbaspirillum
             rubrisubalbicans; Vector: pSport1; Site_1: SalI; Site_2:
             NotI; An unidirectional cDNA library generated from
             [seedlings inoculated with Herbaspirillum
             rubrisubalbicans]. cDNA was prepared from polyA+ mRNA
             using SuperScript Plasmid System Kit (Invitrogen). The
             double-strand cDNAs were fractionated in a sepharose

```

CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:
Pred. No.: 1.28e-24 Length: 585
Score: 522.00 Matches: 128
Percent Similarity: 79.04% Conservatives: 4
Best Local Similarity: 76.65% Mismatches: 17
Query Match: 40.03% Indels: 18
DB: 13 Gaps: 10

US-09-993-808B-2 (1-256) x CA104568 (1-585)

Qy 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaAlaVal 19
Db 113 ATGGGGAAGTACATGCGCAGCAGTGCAGGGCGCGCGGAGGAGTGGCGCGCGTC 172
Qy 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThr---Gly 38
Db 173 GAGGTACAGCAGGTGTCGGGCTCCGACGAGGTCCAGGTCCGCGCGCGGACCGGGGT 232
Qy 39 GlyValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 233 GGTGTGCGAAGTTCGACCGAGGAGGAGGAGCGCTGGCGCGCGCAACGTGGCGGCG 292
Qy 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 293 GGGGAGGCT---GCTGCGCTGCGGCTGAGGGGACGCGGAGCTGCTACATCATCTG 349
Qy 73 ArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValPro 92
Db 350 CGACCGCGATGCTGTTCTATGGCCCGGCTCAGCGCGGAGCGG-----TCGCTCCG 400
Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
Db 401 ACGCCGCGGAGGCTGCT---GGTGCAGCAGCAGGCGCGCGCTCGCGGACGGG 454
Qy 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132
Db 455 CTCTCGGCTGCTCCAGTGGAGTCTNTCGTGACGTCGGG---GGTCTCAACGCGAGG 511
Qy 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
Db 512 AGCCACCCCTGACGCTCC---GAGCTGCGGAGGTTGGCGGGGATCAGCC---CCGAT 565
Qy 153 ValSerAlaAlaSerAsnSer 159
Db 566 GTCTCG---GCGAGCAACTCG 583

RESULT 14
CA104689/c
LOCUS SCJFHR1034G09.1 642 bp mRNA linear EST 23-SEP-2003
DEFINITION SCJFHR1034G09.1 H1 Saccharum officinarum cDNA clone SCJFHR1034G09
3', mRNA sequence.

ACCESSION CA104689

VERSION CA104689.1 GI:34957996

KEYWORDS EST.

SOURCE Saccharum officinarum

ORGANISM Saccharum officinarum

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.

1 (bases 1 to 642)

Vettore, A.L., da Silva, P.R., Kemper, E.L. and Arruda, P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>
Plate: 034 row: G column: 09
Seq primer: SP6 Promoter primer.

FEATURES
source

1..642
Location/Qualifiers
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFHR1034G09"
/lab_host="DH10B"
/clone_lib="HR1"

/note="Organ: seedlings inoculated with Herbaspirillum rubribalbicans; Vector: psporti; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [seedlings inoculated with Herbaspirillum rubribalbicans]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Alignment Scores:
Pred. No.: 2.63e-24 Length: 642
Score: 518.00 Matches: 132
Percent Similarity: 69.12% Conservatives: 18
Best Local Similarity: 60.83% Mismatches: 49
Query Match: 39.72% Indels: 19
DB: 13 Gaps: 10

US-09-993-808B-2 (1-256) x CA104689 (1-642)

Qy 9 ArgGlyAlaAlaGlyAlaGluValAlaAlaValGluValThrGlnValValGlyValArg 28
Db 620 AGGCGCGCTTCGTGGTGGAGGATCCGCGCAATCAGGTACCCAGAT-GTCGGGTCGCT 562
Qy 29 ThrArgSerArgSerAlaAlaAlaThr---GlyGlyValAlaLysValAlaProArgArg 47
Db 561 ACAGATATAGTTTACAGCGCTGCAACTGGGGTGTCTGCGAAGGTTCGACCGGAGG 502
Qy 48 LysArgAla-----ProAlaGlyGluProAlaAlaAlaValSerAla 61
Db 501 AAGAAGCGCTGACACCGCAAGATGACGCGGGGAACCT---ACTGCGATCGGCAAT 445
Qy 62 GlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAlaPro 81
Db 444 GGAGGTACGTGGGAAGCTGCTATCATCATATGCGCAGCGCATGGTGTTCATGTCACG 385
Qy 82 ProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGlyAla 101
Db 384 CATCAGCAGCAG-----CAGTGGCTTCAGACGCGCGGAGCGCTGCT-----GGT 340
Qy 102 AlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSer 121
Db 339 GCAGAACAGCAGGCGCGCGCTAGCCCGCGGACTATCGGTGCTCCAGCAGGAGTCG 280
Qy 122 SerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyrAspAla 141
Db 279 TCGGTGGAGATCGG---GGTTCATCAGCGGAGGAGTACGCTAGCATTCAC---GACACT 226
Qy 142 AlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySer 161
Db 225 CCGGAGGTTGACGGGGATCAGCC---CCGGAAGTTTCG---GCGAGCAGTCGCGGAGC 172
Qy 162 GlyProAspArgGluArgArgGluThrThrProSerSerArgAlaHisGlyLeuLeuSer 181

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 08:37:55 ; Search time 3090 Seconds
(without alignments)
3590.879 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVPLDAGGRFEWAPVVS 256

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09993808/runat_01102004_171702_13026/app.query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09993808 @CGN 1.1 5600 @runat_01102004_171702_13026 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htg.hum:*
40: em.htg.mus:*
41: em.htg.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	669	51.3	1242	6	AX406686 Sequence
2	669	51.3	1266	8	AK103084 Oryza sat
C 3	569	43.6	135670	8	AP005002 Oryza sat
4	399.5	30.6	568	6	AX406682 Sequence
5	279	21.4	681	6	AX406731 Sequence
6	242	18.6	1073	6	AX406733 Sequence
C 7	242	18.6	151208	2	AP003525 Oryza sat
C 8	242	18.6	187916	2	AP003458 Oryza sat
9	240.5	18.4	1045	8	AK064723 Oryza sat
10	238	18.3	1027	8	AK073804 Oryza sat
11	235	18.0	1067	6	AX406720 Sequence
12	215.5	16.5	797	6	AX463079 Sequence
13	212.5	16.3	669	8	ATH301554 Arabidops
14	207	15.9	982	8	LES4411250
15	199.5	15.3	804	6	AX008805 Sequence
16	199.5	15.3	804	6	AX463073 Sequence
17	199.5	15.3	804	8	CRJ002173 Chenopodi
18	190.5	14.6	875	6	A98056 Sequence 3
19	190.5	14.6	875	6	BD075698 Cyclin-de
20	179.5	13.8	861	8	ATH301555 Arabidops
21	178	13.7	870	8	AF123315 Arabidops
22	178	13.7	1290	6	AX463075 Sequence
23	177.5	13.6	1312	8	AK117586 Arabidops
24	170.5	13.1	927	8	AB029483 Pisum sat
25	169.5	13.0	1014	8	LES4411249
26	163	12.5	1072	8	NTA517189
27	159.5	12.2	789	8	NT0297906
28	153	11.7	15311	1	AE007164 Mycobacte
C 29	152.5	11.7	570	8	ATH301556 Arabidops
30	152	11.7	2051	8	AK107643 Oryza sat
31	152	11.7	105659	2	AP003961 Oryza sat
32	152	11.7	162926	2	AP005386 Oryza sat
33	152	11.7	207823	2	AC145542 Oryctolag
34	151.5	11.6	493	6	AX406678 Sequence
C 35	151.5	11.6	178743	10	AL808102 Mouse DNA
36	150.5	11.5	168560	8	AC118132 Oryza sat
37	149.5	11.5	154494	2	AC120885 Oryza sat
C 38	149.5	11.5	316050	1	EX248346 Mycobacte
39	149	11.4	984	6	AX12881 Sequence
40	149	11.4	984	6	AX505332 Sequence
41	148.5	11.4	109081	9	AL358074 Human DNA
C 42	145.5	11.2	7467	8	AY186602 Chlamydom
43	145	11.1	533	6	AX008799 Sequence
44	145	11.1	533	6	AX463068 Sequence
C 45	143.5	11.0	2285	9	BC040188 Homo sapi

ALIGNMENTS

RESULT 1

AX406686
 LOCUS AX406686 1242 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 9 from Patent WO0228893.
 ACCESSION AX406686
 VERSION AX406686.1 GI:21439634
 KEYWORDS Oryza sativa
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
 and Hatzfeld, Y.
 Plant cyclin-dependent kinase inhibitors
 Patent: WO 0228893-A 9 11-APR-2002;
 CROPDISEIGN N V (BE)
 FEATURES
 source Location/Qualifiers
 1..1242
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"
 ORIGIN
 Alignment Scores: 4.6e-22 Length: 1242
 Pred. No.: 669.00 Matches: 161
 Score: 67.42% Conservative: 19
 Percent Similarity: 60.30% Mismatches: 67
 Best Local Similarity: 51.30% Indels: 20
 Query Match: 6 Gaps: 10
 DB:
 US-09-993-808b-2 (1-256) x AX406686 (1-1242)
 Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluAlaAlaValGlu 20
 Db 18 ATGGCAGTACATGAGGAGTTCAGGGGGCCACGGGGAGGAGTGGCGCCATCGG 77
 Qy 21 ValThrGlnValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly--- 39
 Db 78 GTCCAGCAGTGGTGGGTCGCGGACGAGTTCAGGTTCGCGGCGGCGGCGGCGGCG 137
 Qy 40 -----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu 54
 Db 138 AGCAGAAAGTGCAGCGCGCTGCGCGGCTCCACGAGGAGGAGGAGCGCTGCGCG 197
 Qy 55 ProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSer 74
 Db 198 ACGGCGGTCTGGGACTACTCGCGTGCAGCGGCGGAGCTGCTACCTCCAGCTGAGGAC 257
 Qy 75 ArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValProThrPro 94
 Db 258 CGCATGCTGTTTCATGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 314
 Qy 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaAlaGlyLeuSer 114
 Db 315 GCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 371
 Qy 115 ArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlyGlyGlySerHis 134
 Db 372 CGTTGCTCCAGCAGCGCGTCTCGTGGGAC-----GCGGCGGCTCAGGACGAGGCGCTC 425
 Qy 135 ThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
 Db 426 GCGTCCCGCTCC---GACGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 476
 Qy 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrProSerSer 174
 Db 477 TCCGCGAGCGACTCCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 536
 Qy 175 ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAla---GlyHisIleThrGly 193
 Db 537 TTCTCTCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 596

Qy 194 ProSerLeuPro---AlaAlaThrProAlaAlaGluLeuLe---ValPro 208
 Db 597 CGTCCACTACTCTTCTCGGCAACAGCCTCAGCACAGCAGCAGCAGCAGCAGCAGCAG 656
 Qy 209 ProAlaHisGluLeuGlnGluPhePheAlaAlaGluAlaAlaGlnAlaLysArgPhe 228
 Db 657 CCGGCGCGGAGATCGAGGCGGTCTTCTCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAG 716
 Qy 229 AlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPhe 248
 Db 717 GCCGCCAAGTACAACTTCGAGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 773
 Qy 249 GluTrpAlaProValValSer 255
 Db 774 GAGTGGACTCCGTTGGTTCAGC 794
 RESULT 2
 AK103084 1266 bp mRNA linear PLN 24-JUL-2003
 LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J033118G18, full
 DEFINITION insert sequence.
 ACCSSION AK103084 GI:32988293
 VERSION AK103084.1
 KEYWORDS FULL_CDNA; CAP trapper.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 The Rice Full-Length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-Length cDNA Project Team;
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Ohtsuka, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Kojima, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group, Otsu, Y., Murakami, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, T., Itoh, M.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M. and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice
 Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2 (bases 1 to 1266)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, J., Ikeda, R., Ikeda, R., Imamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Matsuyama, T., Miura, J., Miyazaki, A.,
 Masuda, H., Matsubara, K., Murata, M., Nagata, T., Nakamura, R.,
 Namiki, T., Narikawa, R., Niihara, J., Nishi, K., Nomura, K.,
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Ooka, M., Ooka, H.,
 Osato, N., Ota, Y., Otsuka, E., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,

COMMENT

FEATURES

ORIGIN

US-09-993-808B-2 (1-256) x AK703084 (1-1266)

source

FEATURES

22708

```
/clone="P0486G03"

ORIGIN
Alignment Scores:
Pred. No.: 9,08e-16 Length: 135670
Score: 569.00 Matches: 159
Percent Similarity: 46.04% Conservative: 21
Best Local Similarity: 40.66% Mismatches: 68
Query Match: 43.63% Indels: 143
DB: 8 Gaps: 12

US-09-993-808B-2 (1-256) x AP005002 (1-135670)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 80195 ATGGCAAGTACATGAGAGTTACGGGGGGCCACGGGGAGAGTTGGCCCATGAG 80136

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db 80135 GTCAGCAGGTGTTGGCTCCGACGAGGTCGAGTCCGCGAGCGCGCGCGCGAGC 80076

QY 40 -----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu 54
Db 80075 ACGACGAAGTCAAGCGCGCTCGCGCGGTCCACCGAGGAGGAGGCGCTGCTGCCG 80016

QY 55 ProAlaAlaAlaValSerAlaGlyGlyArgGlySerCysTyrTleHisLeuArgSer 74
Db 80015 ACGCGGTGCGGGGACTACTCGCGTACGCGGAGGAGTCTACTCCAGCTGAGGAGC 79956

QY 75 ArgMetLeuPheMetAlaProProGlnProSerValAspSerValProThrPro 94
Db 79955 CGCATGCTGTTTCATGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 79899

QY 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSer 114
Db 79898 GCGGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 79842

QY 115 ArgCysSerSerThrAlaSerSerValAlaSerLeuGlyGlyAlaArgGlySerHis 134
Db 79841 CTTTCTCTCAGCAGCGCGTCTGCTCGTGAC-----GCGCGGCTCAGGACAGGACCTC 79788

QY 135 ThrCysArgSer----- 138
Db 79787 GCGTCCGCTCGACGCTCGCGAGGTGAGGAATCGAACCCCTAGAATTCGAATTCGAAA 79728

QY 139 -----Tyr-AspAlaAl 142
Db 79727 AATGATCCAGATCTTAATCAACCGTTCGTAAGCGTGCCTGCTTTTGGTGTCTTCT 79668

QY 142 aGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerG 162
Db 79667 ACAGGAGGACGAGGAGT-----GTCCCGAGGGCTCCGCGGAGTCTCGCGAGCGG 79614

QY 162 yProAspArgGlu----- 166
Db 79613 CGGTGACCGGAGAGTGGGAGAAATTCCTCCTCATTTCCGCAAAATTCCTGAATTATG 79554

QY 167 -----Ar 167
Db 79553 CGTTCTGACATCCCTCGACGCGGAATCAACCTCGAATTTTCAAAATTTCAAAATTCAC 79494

QY 167 gArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLe 187
Db 79493 GAGAGAAACATTCATCAAGTTTCTCCCGGAGGAGTGGAGGATCTGAGTCTGAGTCT 79434

QY 187 uLa-----GlyHisLysThrGlyProSerLeuPro-----AlaAlaThrProAlaAlaGlu 205
Db 79433 GGTGAGGACAGAGAGCGGAGCGGTCCACTACTCTTCGCGGCAACAGCCTCAGCAGCA 79374

QY 205 ule-----ValProProAlaHisGluThrGlnGluPhePheAlaAlaG 221
Db 79373 AGCCAGCGCGCGGAGAAATTCGCGCGCGCGCGCGGAGTCTTCTCGCGCGCGCGCG 79314

QY 221 uAlaAlaGlnAlaLysArgPheAlaSerLys----- 231
Db 79313 GGAGGCTGAGGCCCAAGCGCTTCGCGCGCAAGTAAAGTATATACAGAAATTTTACCGCTCGC 79254

QY 231 ----- 231
Db 79253 TGTAACATCTAGTAAAGTTACTGCGGCCCTAGTAGCTTCTTCTTCTTCTCGATGC 79194

QY 231 ----- 231
Db 79193 TCGCGGTGGCGTGAAATTTGCGGCTGCTGCGCGGAGCAAGATAATTAACTGCGTGC 79134

QY 232 -----TyrAsnPheAspPheValArgGlyValProLeuAspAlaG 245
Db 79133 TTCTTCCCATCGCGCGCAGGTACAACTTCGAGCTGCTTCGCGCGTGCCTCGACGCC-- 79076

QY 245 yGlyArgPheGluThrAlaProValValSer 255
Db 79075 -GGTCGGTTCAGTGGACTCCGGTGGTCAGC 79046

RESULT 4
AX406682 568 bp DNA linear PAT 14-JUN-2002
LOCUS AX406682 Sequence 5 from Patent WO228893.
DEFINITION AX406682
ACCESSION AX406682
VERSION AX406682.1 GI:21439630
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Drouin, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
and Hatzfeld, Y.
Plant cyclin-dependent kinase inhibitors
Patent: WO 0228893-A 5 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES
Location/Qualifiers
1..568
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Alignment Scores:
Pred. No.: 2,54e-10 Length: 568
Score: 399.50 Matches: 95
Percent Similarity: 81.90% Conservative: 0
Best Local Similarity: 81.90% Mismatches: 14
Query Match: 30.64% Indels: 7
DB: 6 Gaps: 7

US-09-993-808B-2 (1-256) x AX406682 (1-568)

QY 146 GlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArg 165
Db 8 GGGGATCACGTC---CCGATGTCGTCNCGCGAGCAACTCGGGGAGCTCGCGGACCGC 64

QY 166 GluArgArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeu 183
Db 65 GAGAGGAGAGACGACGCGCATCTGTCGAGCCGCGGCGACGCGGCGGAGCTCAGCGATCTG 124

QY 184 GluSerAspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThr 200
Db 125 GAGTGGATCTGGTGGGCGGCGAGAGACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 184

QY 201 ProAlaAlaGluLeuLeuValProProAlaHisGluIleGlnGluPhePheAlaAla 220
Db 185 TCGGCTCGGAGCTGATCTGTCGCGGCGGAGATCCAGGAATTTCTTCGCGCGCGC 244

QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
```


Db 245 GAGCGCGCCATGCAACAGCTTTCTTCCAAAGTACAACTTCGACTCGTCGCGCGCGTG 304

Qy 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
|||||
Db 305 CCCCTCGAGCGC---GGCGGTTGAGTGGAGCGCCAGCGGGTCAAGCATC 349

RESULT 5
AX406731
LOCUS 681 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 54 from Patent WO0228893.
ACCESSION AX406731
VERSION AX406731.1 GI:21439657
KEYWORDS synthetic construct
ORGANISM synthetic construct
SOURCE artificial sequences.
REFERENCE 1
AUTHORS Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
and Hatzfeld, Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 54 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES Location/Qualifiers
source 1..681
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="probe or primer"

ORIGIN
Alignment Scores:
Pred. No.: 7.28e-05 Length: 681
Score: 279.00 Matches: 98
Percent Similarity: 43.51% Conservative: 26
Best Local Similarity: 34.39% Mismatches: 71
Query Match: 21.40% Indels: 90
DB: 6 Gaps: 13

US-09-993-808B-2 (1-256) x AX406731 (1-681)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
|||||
Db 1 ATGGGGAAG---AAGAAGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTG 57

Qy 21 ValThrGlnValValGlyValAlaGlyThrArgSerAlaAlaAlaThrGlyGlyVal 40
|||||
Db 58 GTCGGC-----GGCGTCCGTACCGG----- 78

Qy 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
|||||
Db 79 GCGCGCGTCACGCGGAGGAGGTTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTG 138

Qy 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
|||||
Db 139 CGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 198

Qy 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
|||||
Db 199 CGGAGCAGGAGGCTG----- 213

Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
|||||
Db 214 ---CCCTTCGTGGCGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 270

Qy 113 LeuSerArgCysSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGly 132
|||||
Db 271 GTGGCGGAGCGCGCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 321

Qy 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyAspHisValLeuValAsp 152
|||||
Db 322 -----TGT-----TCGTGGTGGAGGAGGCTATGCGCGCGAG 351

Qy 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
|||||

Db 352 AAGNAGCGCG-----|||||
|||||
Qy 173 SerSerArg-----AlaHisGlyGluLeuSerAspLeuGluSerAsp----- 186
|||||
Db 373 TCGAGCG 432
|||||
Qy 187 -----LeuAlaGlyHisLysThrGlyPro 194
|||||
Db 433 AAGCAGCAAAATGTGCG 489
|||||
Qy 195 SerLeuProAlaAlaThrProAlaAlaGluLeuLeuValProPro-----AlaHisGlu 212
|||||
Db 490 -----GCGAGCAGCAGGAGCTTCAGATGATGATGATGATGATGATGATGATGATG 540
|||||
Qy 213 IleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyr 232
|||||
Db 541 ATCGAGGAGTTCTCGCGCTGCGGAGAGGTTCCGAGGCGCGCGCGCGCGCGCGCGCGCG 600
|||||
Qy 233 AsnPheAspPheValArgGlyValProLeuAspAlaGly-----GlyArgPheGluTrp 250
|||||
Db 601 AACTTCAGCTGTGTGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 660
|||||

Qy 251 AlaProValValSer 255
|||||
Db 661 ACCGCGTGGCGAGC 675

RESULT 6
AX406733
LOCUS 1073 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 56 from Patent WO0228893.
ACCESSION AX406733
VERSION AX406733.1 GI:21439658
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE 1
AUTHORS Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
and Hatzfeld, Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 56 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES Location/Qualifiers
source 1..1073
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Alignment Scores:
Pred. No.: 0.00493 Length: 1073
Score: 242.00 Matches: 117
Percent Similarity: 38.20% Conservative: 27
Best Local Similarity: 31.03% Mismatches: 90
Query Match: 18.56% Indels: 144
DB: 6 Gaps: 19

US-09-993-808B-2 (1-256) x AX406733 (1-1073)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
|||||
Db 1 ATGGGGAAG---AAGAAGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTG 57

Qy 21 ValThrGlnValValGlyValAlaGlyThrArgSerAlaAlaAlaThrGlyGlyVal 40
|||||
Db 58 GTCGGC-----GGCGTCCGTACCGG----- 78

Qy 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
|||||
Db 79 GCGCGCGTCACGCGGAGGAGGTTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTG 138

QY 61 AlaGlyCysArgGlyGlySer-----CysTyrIleHisLeu 72
 Db 139 CGTGGCGTGGCGTGGCGGACGATGGCGAGGCGCGATCTCTCTCGTCTG 198
 QY 73 ArgSerArgMetLeu---PheMetAlaProPro-----GlnProGlnPro 86
 Db 199 CGGACGAGGAGCTGCTTCTGGCGCGCGTGGTGTCTGTCGCGGAGGAGGAGCG 258
 QY 87 SerValAspSerValProThrProValGluAlaAlaAsp----- 99
 Db 259 CTCGGTGAATCGGTG-----CGCGAGCGCGTCTCGTCTGTCGCGCGCGGTG 309
 QY 100 -----GlyAlaAlaGlyGlnGlnGlyAla----- 107
 Db 310 GAATTTGTTGGCTGTCTGGTGGAGGAGGAGGATGCGCGAGAGGTGATGATGAGCC 369
 QY 108 -----AlaLeuAlaAlaGlyLeuSerArg---CysSerSerThrAlaSerSerVal 123
 Db 370 TAGAATTCCTCGCGGCTCGAGTGTCTGATCGCGCGTCTCCATCTCTTGTGATGATGC 429
 QY 124 AsnLeuGlyLeuGly-----GlyGlnArgGlySerHisThrCys 136
 Db 430 GCGTTGGATGTGGTGTGTTGCGAGTTTGCACGCGAGCGAGGATCAG-ACGAGG 488
 QY 137 Arg-----SerTyrAspAlaAlaGluAlaGlyGlyAspHisVal----- 149
 Db 489 AGAGCTCGCTCGCGACTCGCGCTCGCGCGGAGAGGTGATCGAGCTCTCTCCACGCG 548
 QY 150 -----LeuValAspValSerAlaAlaSerAsnSerGlySerGlyPro 163
 Db 549 TTCCTTGTCTGCTTGCATGATTAATACACCGCGTCTCTCAATTGAATATCGCA 608
 QY 164 AspArgGluArgGluThrThrProSerSerArg-----AlaHisGlyGluLeu 180
 Db 609 ATTCAATCAGGAGCGGACGACGCGCTCGAGCCCGCGCGGAGACGCGACTCG 668
 QY 181 SerAspLeuGluSerAsp----- 186
 Db 669 AGCGACGCGGAGTCAACACGAGCGGCAAGCGAATGTGCGCGGAGTTCGACGACC 728
 QY 187 ---LeuAlaGlyHisIysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 205
 Db 729 TCAGCAGCTGCATTTTCACCGGGA-----GCGACGACGAGGAGCTTCAGGATG 776
 QY 206 IleValProPro-----AlaHisGluIleGlnPheAlaAlaAlaAlaAla 223
 Db 777 ATGGCACCGCGCGCGCGCGGAGATCGAGAGTTCCTCGCGCTCGGAGAGGTCC 836
 QY 224 GlnAlaLysArgPheAlaSerLys----- 231
 Db 837 GAGGCGGAGCGCTTCGCCCGCAAGTGTGCTGCATCATATTGTCTCGTCCGTGCTGT 896
 QY 231 ----- 231
 Db 897 GTCGTACATATCGTCTGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 956
 QY 232 -----TyrAsnPheAspPheValArgGlyVal 240
 Db 957 GGAGCTGATTACGTGCGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1016
 QY 241 ProLeuAspAlaGly-----GlyArgPheGluThrAlaProValValSer 255
 Db 1017 CGCTCGACCG 1067

RESULT 7
 AP003525/c 151208 bp DNA linear HTG 21-MAR-2002
 LOCUS Oryza sativa (japonica cultivar-group) chromosome 6 clone P0537F07,
 DEFINITION *** SEQUENCING IN PROGRESS ***.
 ACCESSION AP003525
 VERSION AP003525.1 GI:13676545
 KEYWORDS HTG; HTGS_PHASE2.

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriophytaceae; Oryzoideae; Oryza.

REFERENCE

AUTHORS
 TITLE
 JOURNAL

1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 clone: P0537F07
 Published Only in Database (2001)

2 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 Direct Submission
 Submitted (18-APR-2001) Takuji Sasaki, National Institute of
 Agricultural Resources, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)

NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

LOCATION/Qualifiers
 1..151208
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="6"
 /clone="P0537F07"

FEATURES

source

Alignment Scores:

Pred. No.: 0.414 Length: 151208

Score: 242.00 Matches: 117

Percent Similarity: 38.2% Conservative: 27

Best Local Similarity: 31.0% Mismatches: 90

Query Match: 18.56% Indels: 144

DB: 2 Gaps: 19

US-09-993-808B-2 (1-256) x AP003525 (1-151208)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaAlaValGlu 20

Db 7403 ATGGGGAAG---AAGAAGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7347

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyVal 40

Db 7346 GTCCGC-----GGCGTCCGTACGCG----- 7326

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60

Db 7325 GCGCGCGTCACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7266

QY 61 AlaGlyCysArgGlyGlySer-----CysTyrIleHisLeu 72

Db 7265 CGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 7206

QY 73 ArgSerArgMetLeu---PheMetAlaProPro-----GlnProGlnPro 86

Db 7205 CGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7146

QY 87 SerValAspSerValProThrProValGluAlaAlaAsp----- 99

Db 7145 CTCGGTGAATCGGTG-----GCGGAGCGCGCTTCGTCTGTCGCGCGCGCGGTG 7095

QY 100 -----GlyAlaAlaGlyGlnGlnGlyAla----- 107

```

Db 7094 GAATTGTTGGCTCTCTGTTGAGGAGGAGCTATGGCCGAGAAGTGATTGATGAGCC 7035
Qy 108 -----AlaLeuAlaAlaGlyLeuSerArg---CysSerSerThrAlaSerSerVal 123
Db 7034 TAGAATCTCTCCGGGCTCGAGTCTGATCGCCGCTTCATCTCTTGTGAATGATGC 6975
Qy 124 AsnLeuGlyLeuGly-----GlyGlnArgGlySerHisThrCys 136
Db 6974 GGCTTGGATGTTGTTGAGTTTGCAGGTTGCACCGAGGAGGAGATCAG-ACGAGG 6916
Qy 137 Arg-----SerTyrAspAlaAlaGlyGlyAspHisVal-----149
Db 6915 AGAGCTCGTCCGGGACTCGGCTCGCGCCGAGAGGATGATGAGTCTCTCCACGCG 6856
Qy 150 -----LeuValAspValSerAlaAlaSerAsnSerGlySerGlyPro 163
Db 6855 TTCTGTTGCTTCTGACATGATTAAATACACCGCGCTTCTCAATGAATATATCGCA 6796
Qy 164 AspArgGluArgGluThrThrProSerSerArg-----AlaHisGlyGluLeu 180
Db 6795 ATTCAATCCAGGAGCGGAGCGGCTCGAGCGCGCGCGCGGAGCGGAGTCTG 6736
Qy 181 SerAspLeuGluSerAsp-----186
Db 6735 AGCGAGCGGAGTCAACACGAGGAGGCAAGCAAAATGTCCCGCGGAGTTCGACGACC 6676
Qy 187 ---LeuAlaGlyHisIysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 205
Db 6675 TCAGCAGCTCATTTTCACGGCGA-----CGCAGCAGGAGGCTTCAGGATG 6628
Qy 206 IleValProPro-----AlaHisGluIleGlnIlePheAlaAlaAlaGluAlaAla 223
Db 6627 ATGGCACCGCGCGCGCGGCGGAGAGATCGAGGAGTTCCTCGCGCTCGGAGAGTCC 6568
Qy 224 GlnAlaIysArgPheAlaSerIys-----231
Db 6567 GAGCGCGAGCGCTTCGCGCGCAAGTCTGATGATGATGATGATGATGATGATGATG 6508
Qy 231 -----231
Db 6507 GTCTGATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6448
Qy 232 -----TyrAsnPheAspPheValArgGlyVal 240
Db 6447 GGAGCTGATTAACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6388
Qy 241 ProLeuAspAlaGly-----GlyArgPheGluThrAlaProValValSer 255
Db 6387 CGCTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6337

RESULT 8
AP003458/c
LOCUS 187916 bp DNA linear HTG 31-MAY-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 6 clone P0701E03,
*** SEQUENCING IN PROGRESS ***.
ACCESSION AP003458
VERSION AP003458.2 GI:21280381
KEYWORDS HTG; HTGS, PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
clone:P0701E03
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 187916)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-2001) Takuji Sasaki, National Institute of

```

Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsakak@abrr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)

On May 30, 2002 this sequence version replaced gi:13486944.

NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

source

1. 187916
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="6"
/clone="P0701E03"

ORIGIN

Alignment Scores:

Pred. No.: 0.503 Length: 187916
Score: 242.00 Matches: 117
Percent Similarity: 38.20% Conservative: 27
Best Local Similarity: 31.03% Mismatches: 90
Query Match: 18.56% Indels: 144
DB: 2 Gaps: 19

US-09-993-808B-2 (1-256) x AP003458 (1-187916)

```

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 123090 ATGGGAGG---AAGAGAGAGCGCGCGCGCGCGAGGAGGAGCGCGGGTGTG 123034
Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
Db 123033 GTCCGC-----GGCGTCCGTACGCG-----123013
Qy 41 AlaValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
Db 123012 GCGCGCGTCAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122953
Qy 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
Db 122952 CGTGGCGGTGGCGGTGGCGAGTGGCGAGAGGAGGAGGAGGAGGAGGAGGAGG 122893
Qy 73 ArgSerArgMetLeu---PheMetAlaProPro-----GlnProGlnPro 86
Db 122892 CGAGCAGGAGGAGTGCCTTCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122833
Qy 87 SerValAspSerValProThrProValGluAlaAlaAsp-----99
Db 122832 CTCGGTATTCGGTG-----GCGGAGGCGGCTTCGTCTGTCGTCGCGCGG 122782
Qy 100 -----GlyAlaAlaGlyGlnGlnGlyAla-----107
Db 122781 GAATTGTTGGCTGTTCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 122722
Qy 108 -----AlaLeuAlaAlaGlyLeuSerArg---CysSerSerThrAlaSerVal 123
Db 122721 TAGAATTCCTCCGCGCTCGAGTCTCGATCGCCGCTTCATCTCTGAATGATGC 122662
Qy 124 AsnLeuGlyLeuGly-----GlyGlnArgGlySerHisThrCys 136
Db 122661 GGCCTGGGATGTGTGGTGTTCAGGTTTGCAGGAGGAGGAGGAGGAGGAGGAGGAG 122603
Qy 137 Arg-----SerTyrAspAlaAlaGluAlaGlyGlyAspHisVal-----149
Db 137 Arg-----SerTyrAspAlaAlaGluAlaGlyGlyAspHisVal-----149

```

Db	122602	AGAGCTCCGTCGGCGACTCCGGCTCCGGCCGCGAGAGTGATCGAGCTCTCTCCACGGG	122543		
Qy	150	-----LeuValAspValSerAlaAlaSerAsnSerGlySerGlyPro	163		
Db	122542	TTCTTGCTTTCCTTGACATGATTAATTAACACCGCGTTCCTCAATTAATTAATTCGCA	122483		
Qy	164	AspArgGluArgGluThrThrProSerSerArg-----AlaHisGlyGluLeu	180		
Db	122482	ATTCAATTCAGAGCGGCGACGCGCGTTCGAGCGCGCGGAGACGGGACTCG	122423		
Qy	181	SerAspLeuGluSerAsp-----	186		
Db	122422	AGCGACGCGAGTCAAAACAGGAGGCCAAGCAGCAAAATGTGCGCGCGGAGTTCGACGACC	122363		
Qy	187	---LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu	205		
Db	122362	TCACAGCTTGCATTTCACGCGGGA-----GCGACGACGAGGAGCTTCAGGATG	122315		
Qy	206	IleValProPro-----AlaHisGluIleGlnGluPhePheAlaAlaGluAlaAla	223		
Db	122314	ATGGCACCGCGCGCGGCGGCGACAGATCGAGGAGTTCCTCGCGCTCGGAGAGGTCC	122255		
Qy	224	GlnAlaLysArgPheAlaSerLys-----	231		
Db	122254	GAGGCGGAGCGCTTCGCGCGCAAGTCAGTGCTCATATTCGTCGCGTGTGCT	122195		
Qy	231	-----	231		
Db	122194	GTCGTACATATCGTGTGTGTCTGCTGCTCAAAATCGGCTCGATCGGACATGATGCGCGCATG	122135		
Qy	232	-----TyrAsnPheAspPheValArgGlyVal	240		
Db	122134	GGAGCTGATTAACTGTCGGCTCTCTCTCTCAGGTACAACTTCGACGTGTGTGCGCGCGTG	122075		
Qy	241	ProLeuAspAlaGly-----GlyArgPheGluTrpAlaProValValSer	255		
Db	122074	CGGCTCGACGCGCGCGCGCGCGGCGGCGGTTTCAATGACCGCGGTGGGCGAGC	122024		
RESULT 9	AK064723	1045 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	Oryza sativa (japonica cultivar-group)				cdna clone:002-118-F09, full insert sequence.
DEFINITION	AK064723				
ACCESSION	AK064723.1	GI:32974741			
VERSION	Oryza sativa (japonica cultivar-group)				
KEYWORDS	Oryza sativa (japonica cultivar-group)				
SOURCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
ORGANISM					
REFERENCE	1				
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team., Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Naniki,T., Ohneda,E., Yabagi,W., Suzuki,K., Li,C., Ohtsuka,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawanata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN: Kawai,J., Carrincci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Oka,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.				
	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice				
JOURNAL	Science 301 (5631)	376-379	(2003)		
MEDLINE	22752273				

US-09-993-808B-2 (1-256) x AK064723 (1-1045)	
QY	1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db	88 ATGGGGAAGTACATGCGGAAG-----GGGAAGGTGTGCGGGGAGGTGCGGTGATGAG 141
QY	21 Val---ThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db	142 GTGGGGGGGGCTGCTCGGCTCGGACCGCTCCCGCGCGCTCGCGCTCGCGCGGAGC 201
QY	40 ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
Db	202 ACCTCGTCGCAAGACCGCGGAGAAG-----GGGAGGGGGGACCCCGGTGCGGGCGC 255
QY	60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
Db	256 GCGCGGGGGCGGAG-----TACCTCGAGCTCAGGAGCGCGGCTCGAGAAG 303
QY	80 AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
Db	304 CCGCTCGCACACGCGCG-----GCCAAGGAG 333
QY	100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGlyLeuSerArgCysSerThr 119
Db	334 AAGGAGACCGCGAGGAGGCTTCGCGCGCGCGCGCGCGGTGAGGAT-GCGCGCGCGC 392
QY	120 AlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGly-SerHisThrCysArgSerTy 139
Db	393 GCGCAAGCGCGCGAGGAGTTCGAGCGGAGGTCGAGGTGTCCTTC----- 438
QY	139 rAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSe 159
Db	439 -----GGCGACACGTTCTT-----GACCTCGA 461
QY	159 rGlySerGlyProAspArgGluArgGluThrThrProSerSer-----ArgAlaHi 177
Db	462 CGCGACGCGCCATGAGAGGAGTACCGAGGAGACCAACGCTTCAGTTAAATTAGGAGCTC 521
QY	177 sGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisIleThrGlyProSerLeuPr 197
Db	522 AGAAATGATGAAGCACCCCTGGCTCCCAACT-----AAACCAACACCTTCGATPAC 572
QY	197 oAlaAlaThrProAlaAlaGluLeuIle-----ValProProAlaHisGluIleGl 214
Db	573 TTCCCGCGCGAGATGAGACCTCTGTTGCTGTAGTTCCGAGTTCTCTTGAGATGGA 632
QY	214 nGluPheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPh 234
Db	633 AGAGTTCTTTGCGAGCTGCTGAACACAGCAACATCAGGCTTTCAGAGAGGATATACTT 692
QY	234 eAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrp----- 250
Db	693 CTGTCTGTGACACGACGCGCCACTTCT-----GGACGTGACGATGGACAAGGCTTGA 746
QY	251 -----AlaProValValSer 255
Db	747 CTGCTAGATTTTCATCTTGAGAGCTTCATTGATCTCT 783
RESULT 10	
AK073804	
LOCUS	Oryza sativa (japonica cultivar-group) cDNA clone: J033067M05, full
DEFINITION	insert sequence.
AK073804	
VERSION	AK073804.1 GI:32983827
KEYWORDS	FLU_CDNA; CAP trapper.
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group)
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
	Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of
 Agricultural Sciences Rice Full-Length cDNA Project Team:
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group: Ohtsuki, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Osato, N., Ota, Y.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice
 Science 301 (5631), 376-379 (2003)
 22752273
 12869764

REFERENCE

2 (bases 1 to 1027)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
 Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
 Otsu, N., Ota, Y., Ohtsuki, R., Ryu, R., Saitoh, H., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shingaura, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
 Yoshimura, A.
 Direct Submission

TITLE

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
 Agricultural Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)
 This clone is one of the 28K full-length cDNA clones from japonica
 rice.

COMMENT

URL: <http://cdna01.dna.affrc.go.jp/cdna/>
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
 Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, K., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
 Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
 Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, N., Hiramoto, K.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,

Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
 Yasunishi, A. and Hayashizaki, Y.

FEATURES
 Location/Qualifiers
 1..1027
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J033067M05"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.00716 Length: 1027
 Score: 238.00 Matches: 82
 Percent Similarity: 40.00% Conservative: 22
 Best Local Similarity: 31.54% Mismatches: 76
 Query Match: 18.25% Indels: 80
 DB: 8 Gaps: 11

US-09-993-808B-2 (1-256) x AK073804 (1-1027)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
 DB 130 ATGGCAAGTACATCGCAAGCCCAAGGTGGTCTCCGGCGAGTGGTGGCCGCCGCC 189
 QY 21 ValThrGlnVal-----ValGlyValArgThrArgSerArgSerAlaAla 36
 DB 190 GTCTGGAGCTCGCGCGCGCGCGCTCGGGTGGCGCACCGCGCGCTCCCTCGCGTGT 249
 QY 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
 DB 250 -----CAGAAGAGG----- 258
 QY 57 AlaAlaValSerAlaGlyLysArgGlySerCysTyrIleHisLeuArgSerArgMet 76
 DB 259 -----CAGGGCGGGAG-----TACTCGAGTCTAGGAGCGCAGG 294
 QY 77 LeuPheMetAlaProGlnProGlnProSerValAspSerValProThrProValGlu 96
 DB 295 CTCGAGAAGTCCCTCCCTCCCGCGCG-----CCGCGCGCGAGGAGG 336
 QY 97 AlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCys 116
 DB 337 AGGGCGCGCTCGCGCTCGACTCTGATGCGACGGCGCG----- 378
 QY 117 SerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCys 136
 DB 378 ----- 378
 QY 137 ArgSerTyrAspAlaAlaGluAla-----GlyGlyAspHisValLeuValAsp 152
 DB 379 -----GAGAGCGCGAGGCGGAGGTGTCTTCGGGGGGGAGAACGTCCTC---GAG 426
 QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
 DB 427 CTGGAGGCCATG-----GAAAGGAATACCGAGGAGACGACACCT 465
 QY 173 SerSer-----ArgAlaHisGlyGluLeuSerAspLeuSerAspLeuAlaGlyHis 190
 DB 466 TCGAGCTTGATCAGGAGACCCCGATACGATAGCACCCCTGGATCTACCAAGGCGCAGC 525
 QY 191 LysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuLeuValProProAla 210
 DB 526 CACTCGAGTTCATTGCAAGGTGCAACACCCCGTGGCGCACACACATTAATTCAGCATCA 585
 QY 211 HisGluLeuGlnPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSer 230
 DB 586 GCAGAGCTGGAGCGCTTCCTCGCTGCGGAGAGCAACGCGACGACGAGGTTTCATCGAC 645
 QY 231 LysTyrAsnPheAspPheValArgGlyValProLeuAlaGlyAlaGlyArgPheGluTrp 250

646 AAGTATAACTTTGATCTCTGTGAATGACTGCGCTCTTCCC-----GCCCGGTTTGAATGG 699

Db 646 AAGTATAACTTTGATCTCTGTGAATGACTGCGCTCTTCCC-----GCCCGGTTTGAATGG 699

RESULT 11
 AX406720 LOCUS 1067 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 43 from Patent WO0228893.
 AX406720 ACCESSION
 AX406720.1 GI:21439647
 KEYWORDS
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE
 1 Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
 and Hatzfeld, Y.
 Plant cyclin-dependent kinase inhibitors
 Patent: WO 0228893-A 43 11-APR-2002;
 CROPDESIGN N V (BE)
 JOURNAL
 FEATURES
 Location/Qualifiers
 1..1067
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0101 Length: 1067
 Score: 235.00 Matches: 82
 Percent Similarity: 41.15% Conservative: 25
 Best Local Similarity: 31.54% Mismatches: 73
 Query Match: 18.02% Indels: 80
 DB: 6 Gaps: 12

US-09-993-808B-2 (1-256) x AX406720 (1-1067)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
 DB 39 ATGGCAAGTACATCGCAAGCCCAAGGTGGTCTCCGGCGAGTGGTGGCCGCCGCC 98
 QY 21 ValThrGlnVal-----ValGlyValArgThrArgSerArgSerAlaAla 36
 DB 99 GTCTGGAGCTCGCGCGCGCGCTCGGGTGGCGCACCGCGCGCTCCCTCGCGTGT 158
 QY 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
 DB 159 -----CAGAAGAGG----- 167
 QY 57 AlaAlaValSerAlaGlyLysArgGlySerCysTyrIleHisLeuArgSerArgMet 76
 DB 168 -----CAGGGCGGGAG-----TACTCGAGTCTAGGAGCGCAGG 203
 QY 77 LeuPheMetAlaProGlnProGlnProSerValAspSerValProThrProValGlu 96
 DB 204 CTCGAGAAGTCCCTCCCTCCCGCGCG-----CCGCGCGCGAGGAGG 245
 QY 97 AlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCys 116
 DB 246 AGGGCGAGCTCGCGCT-----GCGACTGCT 272
 QY 117 SerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCys 136
 DB 273 GATGCGACGCGACG----- 287
 QY 137 ArgSerTyrAspAlaAlaGluAla-----GlyGlyAspHisValLeuValAsp 152
 DB 288 -----GAGAGCGCGAGGAGGAGTGTCTTCGGGGGGGAGAACGTCCTC---GAG 335
 QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
 DB 336 CTGGAGGCCATG-----GAAAGGAATACCGAGGAGACGACACCT 374

QY 173 SerSer-----ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHis 190
 Db |||||
 375 TGCAGCTTGATCAGGAGCCCGATAGATTAGCACCCTGGATCTACACAGGCGCAGC 434
 QY 191 LysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuValProAla 210
 Db |||||
 435 CACTCGAGTTCTCATTTGCAAGGTGCAAAACACCGGTGCGCCACAACTATTCCAGCATCA 494
 QY 211 HisGluLeuGlnGluPhePheAlaAlaGluAlaGlnAlaLysArgPheAlaSer 230
 Db |||||
 495 GCAGAGCTGGAGAGCGTTCTTCGCGCCGAGAGCAACCGCAACGACGCGTTTCATCGAC 554
 QY 231 LysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrp 250
 Db |||||
 555 AAGTATAACTTTGATCTGATGACTGCCCTCTTCCC-----GGCCGRTTGAATGG 608

RESULT 12
 LOCUS AX463079 797 bp DNA linear PAT 15-JUL-2002
 DEFINITION Sequence 20 from Patent WO0250292.
 ACCESSION AX463079
 VERSION AX463079.1 GI:21886085
 KEYWORDS Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1 Powke, L.C., Wang, H. and Zhou, Y.
 Modulation of plant cyclin-dependent kinase inhibitor activity
 Patent: WO 0250292-A 20 27-JUN-2002;
 University of Saskatchewan Technologies (CA); Minister of Agriculture and AgriFood Canada (CA)

FEATURES
 source
 1. .797
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0578 Length: 797
 Score: 215.50 Matches: 79
 Percent Similarity: 42.32% Conservative: 34
 Best Local Similarity: 29.59% Mismatches: 93
 Query Match: 16.53% Indels: 61
 DB: 6 Gaps: 12

US-09-993-808B-2 (1-256) x AX463079 (1-797)

QY 1 MetGlyIstYrMetArgIstYsCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
 Db |||||
 10 ATGGGAATACATGAAATAATCAAG-----ATAACTGGCATATCAGCGTCATCGAA 63
 QY 21 ValThrGlnValVal-----GlyValArgThrArgSerArgSerAlaAlaAla 36
 Db |||||
 64 GTCTCTAAGCACACAGCTCCAGTCCAGGTGTTGCAACCCAGAGCGCGTAAACCCCTAGCC 123
 QY 37 ThrGlyGlyValAlaLysValAlaProArgArgGlyAlaProAlaGlyGluProAla 56
 Db |||||
 124 TTG-----AAGCGGCTT-----AATTCCTCC 144
 QY 57 AlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisIeuArgSerArgMet 76
 Db |||||
 145 GCCGCTGATTACGCTTACCTAACAGACTCTCTCTGTCTATCTTCAGCTCCGTCAGCGCGT 204
 QY 77 Leu-----PheMetAlaProGlnProGlnProSerValAspSerVal 91
 Db |||||
 205 CTCGAGAAACCCCTCTTCGCTGATTGACCGAAGACCGG---CCGAGAGTTCAC----- 255
 QY 92 ProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAla 111

Db 256 -----|||
 QY 112 GlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArg 131
 Db |||||
 280 TCCAGGTCTCGGTTGACTCGTTAACTCG----- 309
 QY 132 GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyAspHisValLeuVal 151
 Db |||||
 310 ---GCTCTCTGTAGTCTCAGAGCTCTAATGAGATGAATGTTTTCACATTTCTGTGAGTGT 366
 QY 152 AspValSerAlaAlaSerAsnSer---GlySerGlyProAspArgGluArgGluThr 170
 Db |||||
 367 CAAGTTTCTGTGTGTAACACAGTCTCGGTTTGAATCAACACACACAGCAGGAGAGC 426
 QY 171 ThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHis 190
 Db |||||
 427 ACGCTTGTAAAC-----TTGTTCAGGATATGGAGATCATGTTACACACGGG 474
 QY 191 LysThrGlyProSerLeuProAlaAlaThrPro-----AlaAlaGluLeu 206
 Db |||||
 475 TCTAGCACGAGGTGATGTGCAGAGCAACCAAGAGTACACAGGGAACAAGATAACGTG 534
 QY 207 ValProProAlaHisGluLeuGlnGluPhePheAlaAlaGluAlaAlaGluAlaLys 226
 Db |||||
 535 ATCCGACCACTAGTGAATGGAGGAGTTCTTTCATATGCAGACAGCAGCAGAGG 594
 QY 227 ArgPheAlaSerIstYrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGly 246
 Db |||||
 595 CTATTTCATGAGAGTACAACTTCGACATTTGATGATATCCCTC-----AGCGGA 648

RESULT 13
 LOCUS ATH301554 669 bp mRNA linear PLN 26-JUL-2001
 DEFINITION Arabidopsis thaliana mRNA for cyclin-dependent kinase inhibitor 3 (krp3 gene).
 ACCESSION AJ301554
 VERSION AJ301554.1 GI:14422288
 KEYWORDS cyclin-dependent kinase inhibitor; krp3 gene.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1 De Veylder, L., Beekman, T., Beeckman, T., Maes, S., Naudts, M. and Inze, D.
 Functional analysis of cyclin-dependent kinase inhibitors of Arabidopsis
 Plant Cell 13 (7), 1653-1668 (2001)

AUTHORS
 De Veylder, L., Beekman, T., Beeckman, T., Maes, S., Naudts, M. and Inze, D.
 Submitted (31-OCT-2000) De Veylder L., Departement Plantegenetica, Vlaams Interuniversitair Instituut voor Biologie (VIB), K.U.L.
 Ledeganckstraat 35, B-9000 Gent, BELGIUM

FEATURES
 Location/Qualifiers
 source
 1. .669
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 gene
 1. .669
 /gene="krp3"
 CDS
 1. .669
 /gene="krp3"
 /function="cell cycle control"
 /codon_start=1

/product="cyclin-dependent kinase inhibitor 3"

/protein_id="CAC41617.1"

/db_xref="GI:14422289"

/db_xref="GOA:O9FKB5"

/db_xref="SPTREMBL:O9FKB5"

/translation="MGKMKSKITGDISVMEVKATAPSPGVTRAAKTIALKRLNS
SAADSLPNDSKCVLQSRRLKPSLIEPKOPRVRHSGIKESGSRSDVNSVSP
VAQSNDEDEDFNFSVQVSGCENSLGFESHSTRETPCNFVDMELMTVPQSGSTRS
MCRATKEYTREDDNVIPTTSEMEBEFFAYAEQQQRLPMEKYNFIVNDIFLSGYEWV
QVKP"

ORIGIN

Alignment Scores:

Pred. No.: 0.0673 Length: 669
Score: 212.50 Matches: 79
Percent Similarity: 41.95% Conservative: 33
Best Local Similarity: 29.59% Mismatches: 94
Query Match: 16.30% Indels: 61
DB: 8 Gaps: 12

US-09-993-808B-2 (1-256) x ATH301554 (1-669)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
DB 1 ATGGGGAATACATCAAGAAATCAAG-----ATAACTGGCGATATCAGCGCTCATGGAA 54
QY 21 ValThrGlnValVal-----GlyValArgThrArgSerArgSerAlaAlaAla 36
DB 55 GTCTCTAAAGCAACAGCTCCAGTCCAGGTGTTCGAACCAAGCGCGCTAAACCTTAGCC 114
QY 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
DB 115 TTG-----AAGCGGCTT-----AATTCCTCC 135
QY 57 AlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMet 76
DB 136 GCGCTGATTACGCTTCACTAACGACTCTTCTGCTATCTTCACTGCTGCTGCTGCTGCT 195
QY 77 Leu-----PheMetAlaProGlnProGlnProSerValAspSerVal 91
DB 196 CTCGAGAAACCTCTTCGCTGATTGAACCGAAGCAAGCGCG---CCGAGAGTTTCAAC 246
QY 92 ProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAla 111
DB 247 -----AGATCGGAATTAAGAGTCTGCT 270
QY 112 GlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArg 131
DB 271 TCCAGGTCTCGCGTTGACTCGGTAACTCG----- 300
QY 132 GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuVal 151
DB 301 ---GTTCTGTAGCTCAGAGCTCTAATGAAGATGAATGTTTGCACAAATTCGAGAGTGC 357
QY 152 AspValSerAlaAlaSerAsnSer---GlySerGlyProAspArgGluArgGluThr 170
DB 358 CRAAGTTCTTGCTGTAACACAGTCTCGGTTTGTGATCAAGACACACAGCAGGAGAGC 417
QY 171 ThrProSerSerArgAlaHisGlyGluLeuSerAspLeuSerAspLeuAlaGlyHis 190
DB 418 AGCCCTTGTAAAC-----TTTGTGAGGATATGAGATCATGCTGTTACACAGGAG 465
QY 191 LysThrGlyProSerLeuProAlaAlaThrPro-----AlaAlaGluLeuIle 206
DB 466 TTAGCAGCAGGCTGATGTCAGAGCAACAAAGAGTACACAGGGAACAGATCAAGT 525
QY 207 ValProProAlaHisGluIleGlnGluPheAlaAlaAlaGluAlaAlaGlnAlaLys 226
DB 526 ATCCCGACCACTAGTGAATGAGGAGGATCTTTGTCATATGCAGCAGCAGCAGCAGG 585
QY 227 ArgPheAlaSerLysTyrAsnPheAspValArgGlyValProLeuAspAlaGlyGly 246
DB 586 CTATTCATGGAAGTACAACTTCGACATTTGTAATGATATATCCCTTC-----ACGCGA 639

QY 247 ArgPheGluTrpAlaProVal 253
DB 640 CGTTACGATGGTGGTCAAGTGC 660
RESULT 14
LESA41250 962 bp mRNA linear PIN 18-JUN-2003
LOCUS Lycopersicon esculentum mRNA for p27KIP1-related-protein 2 (krip2)
DEFINITION Lycopersicon esculentum
ACCESSION AJ441250
VERSION AJ441250.2 GI:31980328
KEYWORDS cyclin-dependent kinase inhibitor; krip2 gene;
p27KIP1-related-protein 2.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 Bisbis,B., Joubes,J., Hernould,M., Inze,D., Raymond,P. and
Chevalier,C.
Cyclin-dependent kinase inhibitors are highly expressed in
endoreduplicating tissues during the early development of tomato
fruits
Unpublished
2 Chevalier,C.
Direct Submission
Submitted (09-APR-2002) Chevalier C., Physiologie et biotechnologie
vegetales, INRA-Bordeaux, IBVM, 71 Avenue Bourdeaux, Villenave
d'ornon, 33883, FRANCE
3 (bases 1 to 962)
Chevalier,C.
Direct Submission
Submitted (18-JUN-2003) Chevalier C., Physiologie et Biotechnologie
Vegetales, INRA-Bordeaux, IBVM, 71 Avenue Bourdeaux, Villenave
d'ornon, 33883, FRANCE
On Jun 18, 2003 this sequence version replaced gi:23899380.
COMMENT Location/Qualifiers
FEATURES
1..962
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Wval106"
/db_xref="taxon:4081"
/tissue_type="locular (gel) tissue"
/dev_stage="20 DAA"
1..962
/gene="krip2"
9..566
/gene="krip2"
/function="cyclin-dependent kinase inhibitor"
/codon_start=1
/product="p27KIP1-related-protein 2"
/protein_id="CAD29649.1"
/db_xref="GI:23899381"
/db_xref="SPTREMBL:Q8GT28"
/translation="MGKYLKSGHVMSSSLGVTRSTTTLALQRLSSSSSTPPPSLP
SADSCYLQRLSRRLHKPTPIPCNSHPHSASVDIEISPDNNLHFOHTRSTRESTP
CSLREVDENVNPGSATQTCLNTTQRRNFILNIPSAHEIDFETFAEQQQQRLP
MDKYNFDVVNDVPLSGRYEWINH"

Alignment Scores:
Pred. No.: 0.164 Length: 962
Score: 207.00 Matches: 70
Percent Similarity: 38.74% Conservative: 28
Best Local Similarity: 27.67% Mismatches: 85
Query Match: 15.87% Indels: 70
DB: 8 Gaps: 8
US-09-993-808B-2 (1-256) x LES441250 (1-962)

QY 1 MetGlyIysTyrMetArgLysCysArgGlyAlaAlaGluValAlaValGlu 20
 Db 9 ATGGAAATACTTCGAGAAAGT-----GGCCATGTTATG 44

QY 21 ValThrGlnValGlyValArgThrArgSerArgSerAlaAlaThrGlyVal 40
 Db 45 GAGTCTTCTCTTGGTGGTTCGTACTAGATCGACACCTTAGCA----- 89

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
 Db 90 -----CTTCAACGCTTCAGTCCTTCTTCTTCAACTCCACCCCTCACTGCC 140

QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
 Db 141 TCTGCTTCTGAT-----TCTGTACCTTCAACTTCTGTCCTCCGCGCTTCAC-----AAA 191

QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
 Db 192 CCACCCACCCCATTTCCCTGCCCTTAACCTCACACCCACAC----- 230

QY 101 AlaAlaGlyGlnGlnGlyAlaAlaAlaAlaGlyLeuSerArgCysSerSerThrAla 120
 Db 231 -----TCTGCTTCTGTT 242

QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
 Db 243 GATGAATCTCATTTCTCGCAATAATCTTCAATTTCAACACACTCATAGAGC----- 296

QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
 Db 297 ACTAGAGAACACACCTTTCAGTCTGTGAGGAGGTGGACGAAATGGTCAACCTGGT 356

QY 161 SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
 Db 357 TCTCCGACAGGCAACACAGACTTAACACACAACTCAAGAGAGAGAACTTTATCTCT 416

QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
 Db 417 AGAAAT----- 422

QY 201 ProAlaAlaGluLeuLeuValProAlaHisGluLeuGlnGluPheAlaAlaAla 220
 Db 423 -----ATCCCTTCAGCTCATGAAATCGAGGACTTTTTCACCTTTGCT 464

QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
 Db 465 GACGAGCAACACACCCCTTTTATGGACAAGTACAACTTGTATGGTAAATGATGTA 524

QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProVal 253
 Db 525 CCGCTT-----TCTGSCCGTTATGATGATGATGATGATG 557

RESULT 15
 AX008805
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Chenopodium rubrum (red goosefoot)
 Chenopodium rubrum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Amaranthaceae; Chenopodium.

REFERENCE
 1
 Powke, L.C., Wang, H. and Crosby, W.L.
 Cyclic-dependent kinase inhibitors as plant growth regulators
 Patent: WO 9964599-A 15 16-DEC-1999;
 FOWKE LARRY C (CA); WANG HONG (CA); AGRICULTURE AND AGRIFOOD CANAD
 (CA); CANADA NAT RES COUNCIL (CA); CROSBY WILLIAM L (CA); UNIV
 SASKATCHEWAN TECHNOLOGIES (CA)
 Location/Qualifiers

FEATURES

source 1. 804
 /organism="Chenopodium rubrum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3560"

ORIGIN

Alignment Scores:
 Pred. No.: 0.302 Length: 804
 Score: 199.50 Matches: 67
 Percent Similarity: 42.24% Conservative: 31
 Best Local Similarity: 28.88% Mismatches: 108
 Query Match: 15.30% Indels: 26
 DB: 6 Gaps: 7

US-09-993-808B-2 (1-256) x AX008805 (1-804)

QY 24 ValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGly-ValAlaLysVa 43
 Db 18 GTGGTAGGAGTTAAACCCAGAGCTCGAGACTGCCCTAGCTATGGCGCAGCTGCT----- 72

QY 43 lAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyG 63
 Db 73 -----ACTCCAACCTTCTCTCCGCGAAGATCAAGAAGGTTTC 113

QY 63 YAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAlaProG 83
 Db 114 GAAGTCGTGTATATATTCCTCACTAAGAAAGTCGCGAAGAAATTTGTCGCGCGCGGA 173

QY 83 nProGlnProSerValProThrProValGluAlaAlaAlaAspGlyAlaAlaG 103
 Db 174 GAATTCGCGCAATTAGAAACG-----ACGCCCTTGGAGTTGCGGCGGTGTTGAGGA 227

QY 103 YGlnGlnGlyAlaAlaLeuAlaGlyLeuSerArgCysSerSerThrAlaSerSerVa 123
 Db 228 GGAAGAG-----GTTGCAATTTGCTCGAGT-----AGCGAGGT 260

QY 123 lAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSer--TyrAspAlaAl 142
 Db 261 AATTACTACAGTAGTGGGATTTTCGCGCTCTTGTGCTCAAGCAATTTATGATCAGTT 320

QY 142 aGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerG 162
 Db 321 GAGTTCTAGCGAGCCAGCAAGTAGTTAAGGATGATGATGTTTGGAAATCGTACAGCAGA 380

QY 162 YProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAs 182
 Db 381 TCCAGAGTTGAGAGTGTGAGCGCTGCTCAAGCAAAAG-----GAGAGCCATAG 431

QY 182 pLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAl 202
 Db 432 AACGAAAGCCGAGAGAGCTCAAAATTAGAGCCAGGATTTCCGCGCAGCAAAATCAAC 491

QY 202 aAlaGluLeuLeuValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAl 222
 Db 492 GGTACAGATCAAGATGCGGTCTGATTGAGAAATCGAAGANTCTTCTGCTGCTGAAAA 551

QY 222 aAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLe 242
 Db 552 AGATCTCCAGAAACGCTTCAGCGAAAAGTACAATTTGACATAGATTAAAGACGCTGCCACT 611

QY 242 uAspAlaGlyGlyArgPheGluTrpAlaProVal 253
 Db 612 GAAA-----GGTCGTATGATTTGGGTTCCAATA 639

Search completed: October 2, 2004, 11:11:36

Job time : 3207 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 08:35:39 ; Search time 366 Seconds
(without alignments)
2971.417 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 1304

Sequence: 1 MGKYMKRCGAGAAEVAABE.....VRGVELDAGGFEWAPWWSI 256

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DBV=xlp
-Q=/cgn2_1/USPIO.spool_p/US09993808/runat_01102004_171701_13019/app.query.fasta_1.455
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09993808 @CNC 1.1 708 @runat_01102004_171701_13019 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1304	100.0	1372	7	ABV74603 Maize CKI
2	669	51.3	1242	6	ABK93958 DNA encod
3	664	50.9	841	7	ABV74605
4	612	46.9	723	3	AAA95281
5	413	31.7	572	3	AAN02390 Corn cycl
6	413	31.7	572	3	AAA95276 Corn cycl
7	411	31.5	639	3	AAA95280 Corn cycl
8	399.5	30.6	568	6	ABK93954 Corn zm1c

9	279	21.4	681	6	ABK93981	Abk93981 Full leng
10	242	18.6	1073	6	ABK93982	Abk93982 Rice OsIC
11	235	18.0	1067	6	ABK93971	Abk93971 Full leng
12	219	16.8	1169	3	AAA95292	AAA95292 Wheat cyc
13	215.5	16.5	797	6	AAD40771	Aad40771 Arabidops
14	204	15.6	1005	3	AAA95283	AAA95283 Corn cycl
15	200	15.3	1089	7	ABV74604	ABV74604 Maize CKI
16	199.5	15.3	804	3	AAZ29420	AAZ29420 Chenopodi
17	199.5	15.3	804	6	AAD40768	Aad40768 Chenopodi
18	190.5	14.6	875	2	AAZ25016	Aaz25016 Arabidops
19	178	13.7	870	3	AAA95293	AAA95293 Arabidops
20	178	13.7	870	3	AAZ46465	Aaz46465 Arabidops
21	178	13.7	1290	6	AAD40763	Aad40763 Arabidops
22	163	12.5	1116	3	AAN02402	Aan02402 Cyclin de
23	163	12.5	1116	3	AAA95288	AAA95288 Soybean c
24	152.5	11.7	110000	4	AAI99683	AAI99683
25	151.5	11.6	493	6	ABK93950	Abk93950 Rice OsIC
26	149	11.4	984	6	ABZ12222	Abz12222 Arabidops
27	145.5	11.2	204	3	AAA95286	AAA95286 Rice cycl
28	145.5	11.2	204	3	AAN02400	Aan02400 Rice cycl
29	145	11.1	533	3	AAZ29419	Aaz29419 Arabidops
30	143.5	11.0	2328	7	ABX56289	Abx56289 Human NOV
31	143.5	11.0	2511	7	ACC42663	Acc42663 Human LP3
32	143.5	11.0	2511	7	ACC42663	Acc42663 Human LP3
33	142.5	10.9	472	3	ACA25471	ACA25471 Prokaryot
34	141.5	10.9	4281	7	ACA25471	ACA25471 Prokaryot
35	141.5	10.9	110000	4	AAI99682	AAI99682
36	141.5	10.9	110000	4	AAI99683	AAI99683
37	140.5	10.8	609	3	AAA95279	AAA95279 Wheat cyc
38	140.5	10.8	110000	4	AAI99682	AAI99682
39	138.5	10.6	3921	7	ACA40351	ACA40351 Prokaryot
40	138.5	10.6	110000	4	AAI99682	AAI99682
41	138.5	10.6	110000	4	AAI99683	AAI99683
42	137	10.5	1143	7	ADA70803	Ada70803 Rice gene
43	135.5	10.4	2337	7	ACA40697	ACA40697 Prokaryot
44	135.5	10.4	110000	4	AAI99682	AAI99682
45	135.5	10.4	110000	4	AAI99683	AAI99683

ALIGNMENTS

RESULT 1

ABV74603

ID ABV74603 standard; DNA; 1372 BP.

XX ABV74603;

DT 21-FEB-2003 (first entry)

XX Maize CKI_B coding sequence.

XX Maize; cyclin-dependent kinase inhibitor; CKI; CKI B; plant; corn;
crop yield; root size; plant growth; tassal size; ear size;
male sterility; endoreduplication; gene; ds.

OS Zea mays.

XX Location/Qualifiers

FT Key 134..904

FT CDS /*tag= a

FT FT /product= "CKI_B"

XX WO200281623-A2.

XX 17-OCT-2002.

XX 06-NOV-2001; 2001WO-US044038.

XX 07-NOV-2000; 2000US-0246349P.

XX (PION-) PIONEER HI-BRED INT INC.

XX (ARIZ-) ARIZONA BOARD OF REGENTS.

XX

Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

WPI; 2003-058511/05.

P-PSDB; ABB98757.

Novel cyclin-dependent kinase polynucleotides and their encoded proteins, involved in cell cycle regulation, and useful for altering cell cycle protein content, cell cycle progression, cell number and composition of plants.

Claim 1; Page 64-65; 69pp; English.

The present sequence is the coding sequence (I) for maize cyclin-dependent kinase inhibitor (CKI). CKI B. (I) is useful for modulating the activity of cyclin-dependent kinase (CDK) in a plant such as a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, oil-seed Brassica and millet. Modulating the activity of CKI, preferably modulating downward is useful for providing differential growth in a plant, especially a positive growth advantage and modulating CDK activity upward is useful for increasing crop yield, root size, plant growth, tassel size and/or ear size. Modulating CKI activity is also useful for conferring male sterility and for improving transformation frequencies by increasing the number of cells in cell division. CDK activity can also be modulated for modulating endoreduplication in the endosperm of corn, sorghum, wheat, rice, barley, and millet, where the promoter used is an endosperm-preferred promoter. The cell numbers are modulated in one or more tissues of a plant, comprising root, seed, tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves, stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm, pericarp, meristematic or leaf cells. (I) is also useful for identifying maize CKI interacting proteins, by adducting the nucleic acid sequence to a second nucleic acid sequence encoding a DNA-binding domain

Sequence 1372 BP; 292 A; 396 C; 442 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	2,446-67	Length:	1372
Score:	1304.00	Matches:	256
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-09-993-808B-2 (1-256) x ABV74603 (1-1372)

QY	1	MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaAlaGluValAlaGlu	20
DB	134	ATGGGGAAGTACATGCGCAAGTGCAGGGCGCGCCGAGGTCGCCCGCGCGAG	193
QY	21	ValThrGlnValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal	40
DB	194	GTTACGAGGTGTCGGCGTCGGACAGGTCCAGGTCCGGCGGGCGAGCCGGGTGTC	253
QY	41	AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer	60
DB	254	GCGAAGTCCGCCGAGGAGGAAGAGCGCGCGGGGAGCTGCTCGCGCGGTGAGC	313
QY	61	AlaGlyGlyAspGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla	80
DB	314	GCTGTGGGGAGCGCGGAAGTCTACATCCACTCGTAGCGGACGTCGTTCATGGCA	373
QY	81	ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly	100
DB	374	CGCCCTCAGCCGAGCGCTGCTGTTACTGGTTCGACCCCGGTGGAGGTGCTGATGC	433
QY	101	AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerThrAla	120
DB	434	GCTCAGGACAGCAGGCGCGCGCTCGCGCGCGGCTCTCGCGGTGCTCCAGCAGCGG	493
QY	121	SerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyrAsp	140
DB	494	TGTCGTGTAACATTGGGCTTTGGGGGGTTCAGCGGGAGCCACCTCGCGCTCCTACGAC	553

QY	141	AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly	160
DB	554	GCTCAGAGGTCGGCGGGATCAGCTCTGCTGGATGTCTCGGGCGGAGCAACTCCGGG	613
QY	161	SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu	180
DB	614	AGCGGCCGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	673
QY	181	SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr	200
DB	674	AGCGATCTGGAGTCGATCTGGCGGGGACAAAGTGGCCGTCGTCACCGCGGCAACG	733
QY	201	ProAlaAlaGluLeuLeuValProProAlaHisGluLeuGlnGluPhePheAlaAla	220
DB	734	CCGGCTCGGAGCTGATCTGTCGCCGCCAGCACACGAGATCCAGGATTTCTTCGCCGCC	793
QY	221	GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal	240
DB	794	GAGCGGCGCCAGGCGCAAGCGCTTTGCTTCCAAAGTACAACTTCGACTTCGTCGCGG	853
QY	241	ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerile	256
DB	854	CCCTCTCAGCGCGCGCGGCTTCGAGTGGCGCGGCTGTCAGCATC	901

RESULT 2

ABK93958
ID ABK93958 standard; DNA; 1242 BP.

XX ABK93958;

XX 27-AUG-2002 (first entry)

XX DNA encoding OsICK2 protein.

XX Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.

XX Oryza sativa.

XX WO200228893-A2.

XX 11-APR-2002.

XX 29-JUN-2001; 2001WO-IB001492.

XX 14-JUL-2000; 2000US-0218471P.

XX 13-OCT-2000; 2000US-0241219P.

XX (CROP-) CROPDESIGN NV.

XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D; Hatzfeld Y;

XX WPI; 2002-471311/50.

XX P-PSDB; ABG65670.

XX Novel plant ICK (inhibitors of Cyclin Dependent Kinases) polypeptide used to screen substrates, drugs or compounds which modulate ICK activity and treat disorders characterized by an insufficient or excessive production of ICK inhibitors.

XX Claim 46; Fig 1; 141pp; English.

XX This invention relates to the DNA and protein sequences of novel isolated ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of the invention may be used for treating disorders characterised by insufficient or excessive production of an ICK inhibitor. The protein of the invention may also be used to screen for naturally-occurring ICK substrates, drugs or compounds which modulate ICK activity, as well as to treat disorders characterised by insufficient or excessive production of ICK protein, forms which have decreased or aberrant activity compared to ICK wild type protein. The present sequence represents a cDNA molecule encoding an inhibitor of cyclin dependent kinase (ICK) protein of the invention

Alignment Scores:

Pred. No.: 6.25e-16 Length: 572
 Score: 413.00 Matches: 100
 Percent Similarity: 79.53% Conservative: 1
 Best Local Similarity: 78.74% Mismatches: 18
 Query Match: 31.67% Indels: 8
 DB: Gaps: 8

US-09-993-808B-2 (1-256) x AAA95276 (1-572)

QY 136 CysArgSerTyrAspAla---AlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
 Db 4 TGGCGCTCCGAGCGTGGCGCTGGCGAGGTTACCGGGATCAGTC---CCGGATGTCGC 60
 QY 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro---Ser 173
 Db 61 ACCGCGAGCAACTCGGGGAGCGTCCCGGACCGCGAGAGGAGAGAGAGCGCCATCTCG 120
 QY 174 SerArgAlaHis---GlyGluLeuSerAspLeuGluSerAspLeuAlaGly---HisLys 191
 Db 121 AGCGGGCGCACCGCGCGAGCTCAGCGATCTGGAGTCGGATCTGGTGGCGCGCAGAAG 180
 QY 192 ThrGly---ProSerLeuProAlaAlaThrProAla---AlaGluLeuValProPro 209
 Db 181 ACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 210 AlaHisGluLeuGlnGluPheAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 229
 Db 241 GCACAGAGATCCAGGAATCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 230 SerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGlu 249
 Db 301 TCCAGTACAACTTCGACTTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 357
 QY 250 TrpAlaProValValSerile 256
 Db 358 TGGACGCCAGGGTCAAGATC 378

RESULT 7

AAA95280

ID AAA95280 standard; cdna; 639 BP.

AC AAA95280;

XX 17-JAN-2001 (first entry)

DE Corn cyclin-dependent kinase inhibitor coding sequence #2.

XX Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;

KW cell growth; herbicide; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT CDS 8..388

FT /*tag= a

FT /product= "CDKI"

FT /partial

XX W0200060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX P-ESDB; AAB26249.

XX

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Alignment Scores:

Pred. No.: 9.11e-16 Length: 639

Score: 411.00 Matches: 100

Percent Similarity: 79.53% Conservative: 1

Best Local Similarity: 78.74% Mismatches: 18

Query Match: 31.52% Indels: 8

DB: Gaps: 8

US-09-993-808B-2 (1-256) x AAA95280 (1-639)

QY 136 CysArgSerTyrAspAla---AlaGluAlaGlyGlyAspHisValLeuValAspValSer 154

Db 11 TGGCGCTCCGAGCGTGGCGCTGGCGAGGTTGACGGGGATCAGTC---CCGGATGTCGC 67

QY 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro---Ser 173

Db 68 ACCGCGAGCAACTCGGGGAGCGTCCCGGACCGCGAGAGGAGAGAGAGCGCCATCTCGTCG 127

QY 174 SerArgAlaHis---GlyGluLeuSerAspLeuGluSerAspLeuAlaGly---HisLys 191

Db 128 AGCGGGCGCACGGCGCGAGCTCAGCGATCTGGAGTCGGATCTGGTGGCGCGCAGAAG 187

QY 192 ThrGly---ProSerLeuProAlaAlaThrProAla---AlaGluLeuValProPro 209

Db 188 ACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247

QY 210 AlaHisGluLeuGlnGluPheAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 229

Db 248 GCACAGGAGATCCAGGAATCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 307

QY 230 SerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGlu 249

Db 308 TCCAGTACAACTTCGACTTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 364

QY 250 TrpAlaProValValSerile 256

Db 365 TGGACGCCAGGGTCAAGATC 385

RESULT 8

ABK93954

ID ABK93954 standard; DNA; 568 BP.

XX

AC ABK93954;

XX 07-AUG-2003 (revised)

DT 27-AUG-2002 (first entry)

XX

DE Corn zmlCK1 EST DNA sequence.

XX Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.

XX Zea mays.

XX W0200228893-A2.

Cyclin dependent kinase inhibitor sequences, useful for identifying herbicides and plant growth inhibitors.

Claim 2; Page 43; 59pp; English.

The present sequence is the coding sequence for the corn cyclin-dependent kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA library for sequences similar to those encoding the CDKI from *Chenopodium rubrum*, *Caenorhabditis elegans* and *Arabidopsis thaliana*. CDKI is involved in the cell cycle, and may promote or inhibit cell division and growth.

The coding sequence and the protein it encodes are useful in the production of transgenic plants which produce increased or decreased amounts of the CDKI protein, in the identification of herbicides, in genetic and physical mapping and in the isolation of the CDKI gene in other organisms

Sequence 639 BP; 169 A; 169 C; 192 G; 109 T; 0 U; 0 Other;


```
XX PD 11-APR-2002.
XX PF
XX PP 29-JUN-2001; 2001WO-IB001492.
XX PR 14-JUL-2000; 2000US-0218471P.
XX PR 13-OCT-2000; 2000US-0241219P.
XX PA (CROP-) CROPEDESIGN NV.
XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
XX PI Hatzfeld Y;
XX DR WPI; 2002-471311/50.
XX DR P-PSDB; ABG65674.
XX PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
XX PT to screen substrates, drugs or compounds which modulate ICK activity and
XX PT treat disorders characterized by an insufficient or excessive production
XX PT of ICK inhibitors.
XX PS Claim 41; Page 125; 141pp; English.
XX CC This invention relates to the DNA and protein sequences of novel isolated
XX CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
XX CC the invention may be used for treating disorders characterized by
XX CC insufficient or excessive production of an ICK inhibitor. The protein of
XX CC the invention may also be used to screen for naturally-occurring ICK
XX CC substrates, drugs or compounds which modulate ICK activity, as well as to
XX CC treat disorders characterized by insufficient or excessive production of
XX CC ICK protein, forms which have decreased or aberrant activity compared to
XX CC ICK wild type protein. The present sequence represents a cDNA molecule
XX CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
XX CC invention. (Updated on 07-AUG-2003 to correct OS field.)
XX SQ Sequence 568 BP; 139 A; 149 C; 177 G; 102 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 3,78e-15 Length: 568
Score: 399.50 Matches: 95
Percent Similarity: 81.90% Conservative: 0
Best Local Similarity: 81.90% Mismatches: 14
Query Match: 30.64% Indels: 7
DB: 6 Gaps: 7

US-09-993-808B-2 (1-256) x ABK93954 (1-568)
QY 146 GlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArg 165
DB 8 GGGGATCACGTC---CCGGATGTCGTCNCCGCGAGCAACTCGGGAGCGTCCCGACCGC 64
QY 166 GluArgArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeu 183
DB 65 GAGAGGAGAGAGACGACGCCATCGTCGAGCGGGCGCAGCGCGGAGCTCAGCGATCTG 124
QY 184 GluSerAspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThr 200
DB 125 GAGTCGGATCTGGTGGGGCGCAGAACACTGGCTGCTGCTCGCCGCGGACACAA 184
QY 201 ProAlaAlaGluLeuLeuValProAlaHisGluLeuGlnGluPhePheAlaAla 220
DB 185 TCGGCTCGGAGCTGATCGTCGCCCGCAGCAGCAGGATCCAGGAATCTTCGCGCGCC 244
QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
DB 245 GAGCGGCCCATGCAACGCTTTGCTTCAAGTACAACTTCGACTTCGTCGCGCGGTG 304
QY 241 ProLeuAspAlaGlyGlyArgGlyPheGluTyrAlaProValValSerIle 256
DB 305 CCCCTCGACGCC---GCCCGGTTCTGAGTGACGCCGAGGGGTTCAGATC 349

RESULT 9
ABK93981
```

```
ID XX ABK93981 standard; cDNA; 681 BP.
AC XX
XX ABK93981;
XX DT 27-AUG-2002 (first entry)
XX DE Full length Rice OsICK5 cDNA sequence.
XX KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
XX OS Oryza sativa.
XX OS WO200228893-A2.
XX FN 11-APR-2002.
XX PD 29-JUN-2001; 2001WO-IB001492.
XX PF 14-JUL-2000; 2000US-0218471P.
XX PR 13-OCT-2000; 2000US-0241219P.
XX PA (CROP-) CROPEDESIGN NV.
XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
XX PI Hatzfeld Y;
XX DR WPI; 2002-471311/50.
XX DR P-PSDB; ABG65692.
XX PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
XX PT to screen substrates, drugs or compounds which modulate ICK activity and
XX PT treat disorders characterized by an insufficient or excessive production
XX PT of ICK inhibitors.
XX PS Claim 46; Page 139; 141pp; English.
XX CC This invention relates to the DNA and protein sequences of novel isolated
XX CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
XX CC the invention may be used for treating disorders characterized by
XX CC insufficient or excessive production of an ICK inhibitor. The protein of
XX CC the invention may also be used to screen for naturally-occurring ICK
XX CC substrates, drugs or compounds which modulate ICK activity, as well as to
XX CC treat disorders characterized by insufficient or excessive production of
XX CC ICK protein, forms which have decreased or aberrant activity compared to
XX CC ICK wild type protein. The present sequence represents a cDNA molecule
XX CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
XX CC invention.
XX SQ Sequence 681 BP; 102 A; 176 C; 307 G; 94 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 4,55e-08 Length: 681
Score: 279.00 Matches: 98
Percent Similarity: 43.51% Conservative: 26
Best Local Similarity: 34.39% Mismatches: 71
Query Match: 21.40% Indels: 90
DB: 6 Gaps: 13

US-09-993-808B-2 (1-256) x ABK93981 (1-681)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyValAlaGlyAlaGluValAlaAlaValGlu 20
DB 1 ATGGGAAG---AAGAAGAGCCGCGCGCGCGAGGAGGAGGAGGCGCGGGTGGT 57
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGlyVal 40
DB 58 GTCGGC-----GGCGTCCGTACGCG----- 78
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
DB 79 GCGCGGTCACGCGGAGGAGGGTGGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 138
QY 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
```

```

Db 139 CGTGGCGTGGCGGTGGCGGACGATGCGGAGGCGGATGCTATCTGCGTCTG 198
Qy 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 199 CGGAGCAGGAGGCTG-----
Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
Db 214 ---CCCTTCGTGGCGCGCGGTGTGTCGCGGAGGAGGAGCGCTCGGTATCG 270
Qy 113 LeuSerArgCysSerSerThraSerValAsnLeuGlyLeuGlyGlnArgGly 132
Db 271 GTGGCGGAGGCGGTCTGTCGTCGTCGTCGCGGCGGTGGAATCTTGCGC----- 321
Qy 133 SerHisThrCysArgSerTyrAspAlaGluAlaGlyGlyAspHisValLeuValAsp 152
Db 322 -----TGT-----TCTGTGAGGAGGAGGCTATGCCCGAG 351
Qy 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
Db 352 AAGNAGCGCG-----ACGACGCG 372
Qy 173 SerSerArg-----AlaHisGlyGluLeuSerAspLeuGluSerAsp----- 186
Db 373 TCGAGCGCGCGCGCGGAGACGCGGACTCGAGCGACGCGGAGTCAACACGAGGCGC 432
Qy 187 -----LeuAlaGlyHisLysThrGlyPro 194
Db 433 AAGCAGCAAAATGTGCGCGGAGTTCGACGACTCGAGCAGCTGCATTTCAACGCGGA--- 489
Qy 195 SerLeuProAlaAlaThrProAlaAlaGluLeuLeuValProPro-----AlaHisGlu 212
Db 490 -----GCGACACGAGGAGCTTCAGGATGATGCGCACCGCGCGCGCGCGGAG 540
Qy 213 IleGlnGluPhePheAlaAlaGluAlaGlnAlaGlnAlaLysArgPheAlaSerLysTyr 232
Db 541 ATCGAGGAGTTCCTCGCGCTGCGGAGAGTCCGAGGCGGAGCGCTTCGCGCCCAAGTAC 600
Qy 233 AsnPheAspPheValArgGlyValProLeuAspAlaGly-----GlyArgPheGluTrp 250
Db 601 AACTTCGACGTGTGCGCGGCGTCCGCTCGACGCGCGCGCGCGCGCGCGGCGGTCGAATGG 660
Qy 251 AlaProValValSer 255
Db 661 ACCGCGTGGGCGC 675
RESULT 10
ID ABK93982
XX ABK93982 standard; DNA; 1073 BP.
XX AC ABK93982;
XX DT 27-AUG-2002 (first entry)
XX DE Rice OsICK5 genomic DNA clone.
XX KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
XX OS Oryza sativa.
XX PN WO200228893-A2.
XX PD 11-APR-2002.
XX PF 29-JUN-2001; 2001WO-IB001492.
XX PR 14-JUL-2000; 2000US-0218471P.
XX PR 13-OCT-2000; 2000US-0241219P.
XX PA (CROP-) CROPDESIGN NV.
XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;

```

```

PI Hatzfeld Y;
XX WPI; 2002-471311/50.
XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX Claim 46; Page 140; 141pp; English.
XX This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention
XX
SQ Sequence 1073 BP; 172 A; 287 C; 417 G; 197 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 1.01e-05 Length: 1073
Score: 242.00 Matches: 117
Percent Similarity: 38.20% Conservatives: 27
Best Local Similarity: 31.03% Mismatches: 90
Query Match: 18.56% Indels: 144
DB: 6 Gaps: 19
US-09-993-808B-2 (1-256) x ABK93982 (1-1073)
Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 1 ATGGGGAAG---AAGAAGAAGCGCGACGCGCGCGCGGAGGAGGAGCGCGGGTGGTG 57
Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 58 GTGCGC-----GGCGTCCGTACGCGG----- 78
Qy 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
Db 79 GCCGCGGTACCGCGGAGGAGGCGTGTGCGAGCGCGGAGGAGGAGGAGGAGGAGGAG 138
Qy 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
Db 139 CGTGGCGGTGGCGGTGGCGAGTGGCGGAGACGATGGCGGAGGCGGATGCTATCTGCTCTG 198
Qy 73 ArgSerArgMetLeu---PheMetAlaProPro-----GlnProGlnPro 86
Db 199 CGGAGCAGGAGGCTGCCCTTCGTGGCGCGCGCGGTGTGCTCGCGGAGGAGGAGGAGG 258
Qy 87 SerValAspSerValProThrProValGluAlaAlaAsp----- 99
Db 259 CTCGGTGATTCGGTG-----CGCGAGGCGCGCTTCGTCGTCGTCGCGCGCGGTG 309
Qy 100 -----GlyAlaAlaGlyGlnGlnGlyAla----- 107
Db 310 GAAATTTGTTGGGCTGTCTCGTGAGGAGGAGGCTATGGCGGAGGAGGAGGAGGAGGAG 369
Qy 108 -----AlaLeuAlaAlaGlyLeuSerArg---CysSerSerThrAlaSerSerVal 123
Db 370 TAGAATTCCTCGCGGCTCGAGTGTCTCGATCGCCCGCTTCCATCTCTTCTGCTAATGATGC 429
Qy 124 AsnLeuGlyLeuGly-----GlyGlnArgGlySerHisThrCys 136
Db 430 GGCCTTGGGATGTGTGTTGAGTTTTCAGGTTTTCAGCGGAGGAGGAGGAGGAGGAGG 488
Qy 137 Arg-----SerTyrAspAlaAlaGluAlaGlyGlyAspHisVal----- 149

```

Db 489 AGAGTCCTCGCGGACTCCGCGTCCGCGCGGAGGCTGATCGAGCTCTCTCCACGCG 548
 Qy 150 -----LeuValAspValSerAlaAlaSerAsnSerGlySerPro 163
 Db 549 TTCCTGCTTCTGATGATTAATTAACACCGCTTCTCAATTGAATTATCCCA 608
 Qy 164 AspArgGluArgGluThrThrProSerSerArg-----AlaHisGlyGluLeu 180
 Db 609 ATTCAATCAGAGCGCGGACGCGCGTCCGCGCGCGCGCGGAGAGCGGACTCG 668
 Qy 181 SerAspLeuGluSerAsp----- 186
 Db 669 AGCCACGCGGAGTCAAAACAGGAGGAGCAAGCAGCAAAATGTCCGCGGAGTTCGACGACC 728
 Qy 187 ---LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 205
 Db 729 TCACGAGCTGATTTACGCGGGA-----GGACGACGAGGAGCTTCAGGATG 776
 Qy 206 IleValProPro-----AlaHisGluLeuGlnGluPhePheAlaAlaGluAlaAla 223
 Db 777 ATGGCACCGCGCGCGGCGGCGGAGAGATCGAGGAGTTCTCTCGCGCTCGGAGAGTCC 836
 Qy 224 GlnAlaLysArgPheAlaSerLys----- 231
 Db 837 GAGCGCGAGCGCTTCGCGCGCAAGTGAAGTGTGTCATCACAATATTGTCGTCGTCGT 896
 Qy 231 ----- 231
 Db 897 GTCGTACATATCTGCTGCTGCTCAAAATCGGCTCGATCCGACATGATGCGCGCATG 956
 Qy 232 -----TyrAsnPheAspPheValArgGlyVal 240
 Db 957 GGAGCTGATTAACTGCGCTCTCTCTCTCAGGTACAACTTCGAGCTGTCGCGGCGGTG 1016
 Qy 241 ProLeuAspAlaGly-----GlyArgPheGluTrpAlaProValValSer 255
 Db 1017 CCGCTCACGCGCGCGCGCGCGGCGGTTCGAATGGACCGCGGTGGGCGAGC 1067
 RESULT 11
 ABK93971
 ID ABK93971 standard; cDNA; 1067 BP.
 XX AC ABK93971;
 XX DT 27-AUG-2002 (first entry)
 XX DE Full length Rice OsICK4 cDNA sequence.
 XX KW Plant; ss; primer; PCR; inhibitor of cyclin dependent kinase; ICK.
 XX OS Oryza sativa.
 XX PN WO200228893-A2.
 XX PD 11-APR-2002.
 XX PF 29-JUN-2001; 2001WO-IB001492.
 XX PR 14-JUL-2000; 2000US-0218471P.
 XX PR 13-OCT-2000; 2000US-0241219P.
 XX PA (CROP-) CROPDESIGN NV.
 XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 XX PI Hatzfeld Y;
 XX PR WPI; 2002-471311/50.
 XX PR P-PSDB; ABG65691.
 XX PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.

XX PS Claim 46; Fig 2; 141pp; English.
 CC This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a PCR primer
 CC specific for an inhibitor of cyclin dependent kinase (ICK) gene used in
 CC the examples of the specification

SQ Sequence 1067 BP; 247 A; 292 C; 300 G; 226 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 2.55e-05 Length: 1067
 Score: 235.00 Matches: 82
 Percent Similarity: 41.15% Conservative: 25
 Best Local Similarity: 31.54% Mismatches: 73
 Query Match: 18.02% Indels: 80
 DB: 6 Gaps: 12

US-09-993-808B-2 (1-256) x ABK93971 (1-1067)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
 Db 39 ATGGGCAAGTACATGCGCAAGGCAAGTGTGTCTCCGCGAGGTGTGTGCGCGCGCC 98
 Qy 21 ValThrGlnVal-----ValGlyValArgThrArgSerArgSerAlaAlaAla 36
 Db 99 GTCATGAGTCCGCGCGCGCGCGCGCTCGGGTGCACCGCGCGCGCTCTCTCGCGCTG 158
 Qy 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
 Db 159 -----CAGAGAGG----- 167
 Qy 57 AlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMet 76
 Db 168 -----CAGGCGCGGAG-----TACCTCGAGCTCAGGAGCGCGAGG 203
 Qy 77 LeuPheMetAlaProGlnProGlnProSerValAspSerValProThrProValGlu 96
 Db 204 CTCGAGAAAGTCT 245
 Qy 97 AlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCys 116
 Db 246 AGGCGGAGCGGTGGGT-----GGAGCTGCT 272
 Qy 117 SerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCys 136
 Db 273 GATCGCAGCGGCGAG----- 287
 Qy 137 ArgSerTyrAspAlaAlaGluAla-----GlyGlyAspHisValLeuValAsp 152
 Db 288 -----GAGAGCGCGGAGCGGAGGTGTCGTTGCGGGGGGAGAACGTCCTC---GAG 335
 Qy 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
 Db 336 CTCGAGGCCCATG-----GAAAGGAATACAGGAGGAGACACCT 374
 Qy 173 SerSer-----ArgAlaHisGlyLeuLeuSerAspLeuSerAspLeuAlaGlyHis 190
 Db 375 TGCAGCTTGATCAGGACCGCGATACGATTAGCACCCCTGGATCTACACACAGCGCAGC 434
 Qy 191 LysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuLeuValProProAla 210
 Db 435 CACTCGAGTTCTCATTCGCAAGGTGCAACACCGCTGCGCCACACCATATTATCCAGCATCA 494
 Qy 211 HisGluIleGlnGluPheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSer 230

495	GCAGAGCTGGAAGCGTCTCTTCGCGCCGAAGAGCAACGCACACGACAGGCTTTTCATCGAC	554
Db		
231	LysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyValArgPheGluTrp	250
Qy		
555	AAGTATAACTTTGATCTCTGTAATGATGCGCTCTTCCC-----GGCGERTTTGAATGG	608
Db		

RESULT 12
AAA95292
ID AAA95292 standard; CDNA; 1169 BP.

Qy	21	ValThrGlnVal	---	ValGlyValArgThrArgSerAlaAlaAlaThrGlyGly	39
Db	195	GTCGCCGCCGCCGCTAGGGGTCCGACCCGCCGACGAGCGCTCCG	---	---	242
Qy	40	ValAlaLysValAlaProArgLysArgAlaProAlaGlyGluProAlaAlaVal	59		
Db	243	---	---	---	272
Qy	60	SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet	79		
Db	273	GCCAAAG	---	---	323
Qy	80	AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp	99		
Db	324	CTGCCCCCGCCGCCGCG	---	---	341
Qy	100	GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThr	119		
Db	342	---	---	---	365
Qy	120	AlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyr	139		
Db	366	GCGAGCGTGTCGAG	---	---	380
Qy	140	AspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSer	159		
Db	381	---	---	---	434
Qy	160	GlySerGlyProAspArgGluArgGlnThrThrProSerSer	177		
Db	435	GCCATGGGG	---	---	488
Qy	178	GlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSer	195		
Db	489	GGACGATAAGC	---	---	533
Qy	196	---	---	---	212
Db	534	AATTCCCATCGCAGGGTGCAAGCTCCAGCGCGCCAT	---	---	590
Qy	213	IleGlnGluPhePheAlaAlaGluAlaGlnAlaLysArgPheAlaSerLysTyr	232		
Db	591	ATGAATGATGTTCTTCTGTCTGCGGAGCAACCGCAACAGCAAGCCTTCATTGACAAAGTAC	650		
Qy	233	AsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrp	250		
Db	651	AACCTTGTATCCTGTGAACGACTGCTCTCTCCA	---	---	698
RESULT 13					
AAD40771					
ID	AAD40771 standard; cDNA; 797 BP.				
XX	AAD40771;				
XX	AC				
DT	30-OCT-2002 (first entry)				
XX	Arabidopsis thaliana ICN7 full-length cDNA.				
DE					
XX					
KW	Plant development; cyclin-dependent kinase; CDK inhibitor; ICN1; ICN2; ICN6; ICN7; ICN8; ICN9; morphogenesis; maturation; enlargement;				
KW	plant breeding; growth; gene; ss.				
XX	Arabidopsis thaliana.				
OS					
XX					
FH	Key				
FT	CDS				
FT	/*tag= a				
FT	/product= "ICN7 full-length protein"				
XX					
PN	WO200250292-A2.				
XX					
XX	27-JUN-2002.				
PD					

CC identification of herbicides, in genetic and physical mapping and in the
 CC isolation of the CDKI gene in other organisms
 XX
 SQ Sequence 1005 BP; 214 A; 315 C; 287 G; 188 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 0.00152 Length: 1005
 Score: 204.00 Matches: 71
 Percent Similarity: 41.67% Conservative: 34
 Best Local Similarity: 28.17% Mismatches: 79
 Query Match: 15.64% Indels: 68
 DB: 3 Gaps: 10

US-09-993-808B-2 (1-256) x AAA95283 (1-1005)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
 Db 175 ATGGCAAGTACATCGCAAGCCCAAG-----GCTTCAGCGAGGTTGTCATCATGGAT 228
 QY 21 ValThrGlnVal---ValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
 Db 229 GTGCGCGCGCTCCGCTCGAGTCCGACCCGAGCGCGCGCTCGCGTG----- 279
 QY 40 ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
 Db 280 -----CAGCGTCTGCGAGGACGACGACGAGCGAGTGGGAGGAA 315
 QY 60 SerAlaGlyGlyAspGlyCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
 Db 316 GGTGCTGCGCGCGAG-----TACCTGGAGCTAAGGACCGGAGGCTCGAGAG 363
 QY 80 AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
 Db 364 CTGCGCGCGCGCGCG----- 378
 QY 100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThr 119
 Db 379 -----GCGACCAAG 387
 QY 120 AlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyr 139
 Db 388 AGGAGGTTCG-----GGCGGAGGAAAGCGCGAGCGCCGCGCAACT 432
 QY 140 AspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSer 159
 Db 433 AAGGAGGCTGAGCGCTGCTACCGGAGAACATGCTCGAGTTGGAGCCATG----- 483
 QY 160 GlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGlu 179
 Db 484 -----GAGAGGATTACAGGAGACGACGCGCTCGAGCTTGATTAACACCCAG 531
 QY 180 LeuSerAspLeuGlu---SerAspLeuAlaGlyHisLysThrGlyProSerLeuProAla 198
 Db 532 ATGACTAGCACTCTGGTTCACGAGATCCGCGCACTTTGGCCACGCGAGGTGACGCT 591
 QY 199 AlaThrProAlaAlaGluLeuValProAlaHisGlnIleGlnIleGlnPheAla 218
 Db 592 CCTCGGTCCAGCC-----GTCCCAAGTCTAGGAGATGATGATGACTTCGT 642
 QY 219 AlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArg 238
 Db 643 GCGCAACAGCGAGCGCAACAGCAGGAGTTTCATTGACAGTACAACTTCGATCTCGCAAC 702
 QY 239 GlyValProLeuAspAlaGlyGlyArgPheGluTrp 250
 Db 703 GACTGCCCTCTCCCA-----GCGAGGTTTGATGG 732

RESULT 15

ID ABV74604
 ID ABV74604 standard; DNA; 1089 BP.

XX ABV74604;

XX

DT 21-FEB-2003 (first entry)

XX Maize CKI_C coding sequence.

XX Maize; cyclin-dependent kinase inhibitor; CKI; CKI_C; plant; corn;
 KW crop yield; root size; plant growth; tassel size; ear size;
 KW male sterility; endoreduplication; gene; ds.

XX Zea mays.

XX Key Location/Qualifiers
 FT 154..726
 FT /*tag= a
 FT /product= "CKI_C"

XX WO200281623-A2.

XX 17-OCT-2002.

XX 06-NOV-2001; 2001WO-US044038.

XX 07-NOV-2000; 2000US-0246349P.

XX (PION-) PIONEER HI-BRED INT INC.
 PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

XX WPI; 2003-058511/05.
 DR P-PSDB; ABB98758.

XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
 PT involved in cell cycle regulation, and useful for altering cell cycle
 PT protein content, cell cycle progression, cell number and composition of
 PT plants.

XX Claim 1; Page 66-67; 69pp; English.

XX The present sequence is the coding sequence (I) for maize cyclin-
 CC dependent kinase inhibitor (CKI), CKI_C. (I) is useful for modulating the
 CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
 CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
 CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
 CC preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain

XX Sequence 1089 BP; 251 A; 326 C; 304 G; 208 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.00282 Length: 1089
 Score: 200.00 Matches: 70
 Percent Similarity: 40.08% Conservative: 35
 Best Local Similarity: 26.72% Mismatches: 69
 Query Match: 15.34% Indels: 88
 DB: 7 Gaps: 11

US-09-993-808B-2 (1-256) x ABV74604 (1-1089)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyValAlaAlaValGlu 20
 XX
 XX
 XX

Db 154 ATGGGCAAGTACATCGCGCAAGGCAAG-----GCTTCCAGCGAGGTGTGTCATCATGGAT 207
QY 21 ValThrGlnVal---ValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db 208 GTCGCCCGCGTCCGCTCGGAGTCCGCACCGCGCGCGCTCGCGCTG----- 258
QY 40 ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
Db 259 -----CAGCGTCTGAGGAGCAGCAGCAGCAGCAGTGGAGGAA 294
QY 60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
Db 295 GGTGCTGCGCGCGAG-----TACCTGAGCTTAGGAACCGAGGCTCGAGAAG 342
QY 80 AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAsp 99
Db 343 CTGCCCGCGCGCGCG----- 357
QY 100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThr 119
Db 358 -----GCGACCAAG 366
QY 120 AlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyr 139
Db 367 AGGAGGTCG-----GGCGGAGGAAGCGCGAGCGCGCGCAACT 411
QY 140 AspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSer 159
Db 412 AAGGAGCTGAGCGCTGTCGCGGAGAAACATGCTCGAGTTGGAGGCATG----- 462
QY 160 GlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGlu 179
Db 463 -----GAGAGGATTACCAGGAGAGCAGCGCTTGACCTTGATTACACCCAG 510
QY 180 LeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAla 199
Db 511 ATGACT-----AGCACTCTCTGGGTCC 531
QY 200 ThrProAlaAla-----GluLeuIleValProProAlaHis----- 211
Db 532 ACGAGATCCAGCCACTCTTGCCACCGAGGTGAACGCTCTCGGTGCGCGCGTCCCA 591
QY 212 -----GluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPhe 228
Db 592 AGTTCGAGGAGATGATGATGACTTCTGCTGCCAAGCAGCGACGCCAACACAGGATTTC 651
QY 229 AlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPhe 248
Db 652 ATTCACAAGTACAACTTCGATCTGCAAGCAAGCACTGCGCTCTCCCA-----GGCAGGTTT 705
QY 249 GluTyr 250
Db 706 GAGTGG 711

Search completed: October 2, 2004, 10:18:03
Job time : 372 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 09:59:54 ; Search time 85 Seconds
(without alignments)
1671.382 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 1304

Sequence: 1 MGKYMRCRGAAGAAVE.....VRGVLDAGGFEWAPVVS 256

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USFIO_spool_P/US0993808/runat_01102004_171703_13045/app_query.fasta_1.455
-DB=issued Patents NA -OFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808 @CIGN 1 1 105 @runat_01102004_171703_13045 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTC.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	152.5	11.7	4403765	3	US-09-103-840A-2
C 2	141.5	10.9	4403765	3	US-09-103-840A-2
C 3	141.5	10.9	4411529	3	US-09-103-840A-1
C 4	140.5	10.8	4411529	3	US-09-103-840A-1
C 5	134	10.3	2561	4	US-09-616-289-48
C 6	130.5	10.0	1845	4	US-09-252-991A-3639
C 7	130.5	10.0	12425	4	US-09-616-289-50
C 8	129.5	9.9	4848	4	US-08-955-957A-1
C 9	129.5	9.9	4848	4	US-08-955-957A-1
C 10	129.5	9.9	4848	4	US-08-955-957A-6
C 11	126.5	9.7	33529	3	US-09-144-085-3
C 12	125.5	9.6	1143	4	US-09-457-568-7

13	125.5	9.6	1143	4	US-09-457-646-7	Sequence 7, Appli
14	125.5	9.6	1143	4	US-08-902-572-5	Sequence 5, Appli
15	125.5	9.6	1143	4	US-09-516-065-7	Sequence 7, Appli
C 16	125	9.6	861	4	US-09-347-613C-8	Sequence 8, Appli
17	125	9.6	2481	4	US-09-894-988A-35	Sequence 35, Appli
18	125	9.6	13987	2	US-08-804-227C-13	Sequence 13, Appli
19	123.5	9.5	807	4	US-09-403-768-1	Sequence 1, Appli
20	123.5	9.5	1983	4	US-09-252-991A-16300	Sequence 16300, A
21	123	9.4	3915	4	US-09-489-039A-6278	Sequence 6278, Ap
C 22	123	9.4	53526	3	US-08-658-136-2	Sequence 2, Appli
C 23	123	9.4	53577	3	US-08-658-136-1	Sequence 1, Appli
24	122.5	9.4	34094	4	US-09-292-034-1	Sequence 1, Appli
C 25	122.5	9.4	43280	2	US-08-804-227C-1	Sequence 1, Appli
C 26	121.5	9.3	1070	4	US-09-470-443-7	Sequence 7, Appli
27	121.5	9.3	1659	3	US-09-083-351-3	Sequence 3, Appli
28	121.5	9.3	1659	3	US-09-083-352-3	Sequence 3, Appli
29	121.5	9.3	3946	3	US-09-083-351-1	Sequence 1, Appli
30	121.5	9.3	3946	3	US-09-083-352-1	Sequence 1, Appli
31	121	9.3	2472	4	US-09-252-991A-8197	Sequence 8197, Ap
32	121	9.3	43280	2	US-08-804-227C-1	Sequence 1, Appli
33	120	9.2	1098	4	US-09-457-568-9	Sequence 9, Appli
34	120	9.2	1098	4	US-09-457-646-9	Sequence 9, Appli
35	120	9.2	1098	4	US-08-902-572-7	Sequence 7, Appli
36	120	9.2	1098	4	US-09-516-065-9	Sequence 9, Appli
C 37	120	9.2	1323	4	US-09-252-991A-3022	Sequence 3022, Ap
C 38	120	9.2	1965	4	US-09-252-991A-9230	Sequence 9230, Ap
C 39	120	9.2	2805	4	US-09-252-991A-9084	Sequence 9084, Ap
40	120	9.2	4190	3	US-08-938-291A-2	Sequence 2, Appli
41	120	9.2	4190	4	US-09-589-619-2	Sequence 2, Appli
42	119.5	9.2	507	4	US-09-105-470B-28	Sequence 28, Appli
C 43	119	9.1	702	4	US-09-252-991A-8108	Sequence 8108, Ap
C 44	119	9.1	1533	4	US-09-252-991A-12896	Sequence 12896, A
C 45	119	9.1	1854	4	US-09-894-844-103	Sequence 103, App

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FRISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 35.4
Score: 4403765
Percent Similarity: 35.91%
Best Local Similarity: 28.96%
Query Match: 11.69%
DB: 3
Length: 4403765
Matches: 75
Conservative: 18
Mismatch: 100
Indels: 66
Gaps: 10

US-09-993-808B-2 (1-256) x US-09-103-840A-2 (1-4403765)


```
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 294 Length: 4411529
Score: 141.50 Matches: 72
Percent Similarity: 38.96% Conservative: 18
Best Local Similarity: 31.17% Mismatches: 100
Query Match: 10.85% Indels: 42
DB: 3 Gaps: 9

US-09-993-808B-2 (1-256) x US-09-103-840A-1 (1-4411529)
QY 6 ArglyscysArgGlyAlaAlaGlyAlaGluValAlaAlaGluValThrGln----- 23
Db 928765 CGCAGAGTGGCGCGGTAGCGTGGCGCC--GTTGCGCGCGTGGCGCGTTAACCCCGAC 928821
QY 24 -----ValValGlyValArgThr---ArgSerArgSerAlaAlaAlaThrGly 38
Db 928822 TCCGGTCCCGCGGTGGCGCGGTGGCCACCTGCCCGCGACGAATCCGAAGCGCTCACCGCC 928881
QY 39 GlyValAlaAlaValAlaProArgArgAlaAlaProAlaGlyGluProAlaAlaAla 58
Db 928882 GGCACCGCC--GCTGGCGCGGTACCAACGAGACCCCGCGCGTGTCCACCGCGGCC 928938
QY 59 ValSerAlaGlyGlyAspGlyGly-SerCysTyrIleHisLeuArgSerArgMetLeuPh 78
Db 928939 GCCACCGCCCGCCAGCAGCGCGGTCCG-----CTGCC 928971
QY 78 eMetAlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAl 98
Db 928972 TCCGGCGCGCGCGTGGCGCGGTGTCGTGGCGCTCCGCGCAACCCCGCC--GACGCGCG 929030
QY 98 aAspGlyAlaAlaGlyGlnGln-----GlyAlaAlaLeuAlaAlaGlyLe 113
Db 929031 CGATCCCGCGCGCATCATCCGAGGCGATCGCGCGCGCGCGCGCGCGCGCGCGCTAC 929090
QY 113 uSerArgCysSerThrAlaSerValAlaSerValAlaGlyGlyGlyGlyGlyGly 133
Db 929091 CAGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 929150
QY 133 rHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGly----- 146
Db 929151 CCAATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 929210
QY 147 -----AspHisValLeuValAspValSerAlaAlaSerAsnSerGlyPro-- 163
Db 929211 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 929270
QY 164 -----AspArgGluArgArgGluThrThrProSerSerArgAlaHisGlyGluLe 180
Db 929271 CGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 929328
QY 180 uSerAspLeuGluSerAsp-LeuAlaGlyHisIysThrGlyProSerLeuProAlaAla 200
Db 929329 -----GCCGAACCCGATGAAGTCGGGGGCACAGCGAAGCGCGCGCGCGCGCGCG 929381
QY 200 hrProAlaAlaGluLeuValProPro 209
Db 929382 CGCGGTCCCGATGAGTGGCGCGCACCA 929410

RESULT 4
US-09-103-840A-1/c

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 356 Length: 4411529
Score: 140.50 Matches: 81
Percent Similarity: 38.46% Conservative: 19
Best Local Similarity: 31.15% Mismatches: 104
Query Match: 10.77% Indels: 57
DB: 3 Gaps: 11

US-09-993-808B-2 (1-256) x US-09-103-840A-1 (1-4411529)
QY 6 ArglyscysArg-GlyAlaAlaGlyAlaGluValAlaAlaGluValThrGlnValva 25
Db 3948721 CGCTTTGGCGCGGTGCCCGCGGTGCCCGGTGGTGGCGCTGTGTAGTCCGT 3948662
QY 25 lGlyVal-----ArgThrArgSerArgSerAlaAlaAlaThrGlyGlyValAlaLysVa 43
Db 3948661 CGGTGTTAGCGCGCTTTCGCCCGCGGTGCCCGGTGCCCGGTGCCCGGTGTAGT 3948602
QY 43 lAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGl 63
Db 3948601 TGGCCCC-----GCCGTACCGCGCGC---CCCGCGCTGCCCGCGCGCGCGG 3948554
QY 63 yAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla-ProProG 83
Db 3948553 TGGCGCGGTGGCGCGCTTTCGCCCGGTGCCCGCGGTGCCCGGTGCCCGGTGCCG 3948494
QY 83 lNProGlnProSerValAspSerValPro-----ThrProValGluAlaA 98
Db 3948493 CCGCGGTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3948435
QY 98 lAlasp-----GlyAlaAlaGlyGlnGlnGlyAlaA 108
Db 3948434 CTGATGCCAGCACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3948375
QY 108 lAlaAlaAlaGlyLeuSerArgCysSerThrAlaSerSer----- 122
Db 3948374 TTGTGCGCGCGCGCACAGCTGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3948315
QY 123 -----ValAsnLeuGlyLeuGlyGlyGlnArgGlys 133
Db 3948314 AACCCGCGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3948255
QY 133 erHisThrCysArgSerTyrAspAlaAlaGlu-----AlaGlyGlyAspHisV 149
Db 3948254 CCGTCACCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT- 3948196
QY 149 alLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArgLysArgG 169
Db 3948195 -----GCCCGCGGTGCCGTCCCGTCCCGTCCCGCGCGCGCGCGCGCGCG 3948168
QY 169 luThrThrPro-----SerSerArgAlaHisGlyGluLeuSerAspLeuSerAspL 187
Db 3948168 -----CGCT 3948168
```


; SEQ ID NO 7
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-568-7

Alignment Scores:

Pred. No.: 0.0983 Length: 1143
Score: 125.50 Matches: 80
Percent Similarity: 37.01% Conservative: 24
Best Local Similarity: 28.47% Mismatches: 88
Query Match: 9.62% Indels: 90
DB: 4 Gaps: 16

US-09-993-808B-2 (1-256) x US-09-457-568-7 (1-1143)

```
QY 9 ArgGlyAlaAlaGlyAlaGlu-----ValAlaAlaValGluValThrGlnValValGly 26
Db 105 AGAGGAGGTGCGGCGCTGCTGGAGCGGGCGCTGCCAACGCCACCGA-ATAGTTACG 163
QY 27 ValArgThrArgSerArgSer-----AlaAla-AlaThrGlyGlyValAlaLysVa 43
Db 164 GTGGAGGCGATCCAGTTCATGATGATGAGCGCGCGCGAGTGGCGAGCTGCT--GC 220
QY 43 lAlaProArgArg-----Ly 48
Db 221 TGCTCCACGGCGGAGGCCAACCTGCGCGCCCGCCACTCTCACCCGACCGGTGCACG 280
QY 48 sArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlySerCy 68
Db 281 ACGTGTCGCGGAGGGCTTCT-----GGACACGCTGGTGTGCT--GC 320
QY 68 sTyrIleHisLeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerVa 88
Db 321 -----GCACCGGGCGCG----- 332
QY 88 lAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGln----- 104
Db 333 -----GGCGCGCTGGACGTGCGCGCGTGCCT 358
QY 105 -----GlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAl 120
Db 359 GGGGCGCTGCGCGGTGGAGTCCGTCGAGGAGTGGCCATCGCATGTGCGACGGTACC 418
QY 120 aSerSerValAsnLeuGlyGlyGln-ArgGlySerHisThrCysArgSerTyrA 140
Db 419 TGGCGC-----GGCTGCGGGGGGCGCACGAGCGAGTAAACCATGCCCGC--ATAG 466
QY 140 spAlaAlaGluAlaGlyGlyAspHisValLeuValValSerAlaAlaSer---AsnS 159
Db 467 ATCCCGGAGAGTCCCTCAGAC-----ATCCCGATGTGGGGGGGCGAGCGGGGTG 520
QY 159 erGlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyG 179
Db 731 TCGATTTTCAGATCACAAACCCCTAGAG-----GGCAAGTACGAGTGGCAAGAGTG 783
```

RESULT 13

US-09-457-646-7
; Sequence 7, Application US/09457646
; Patent No. 6420345
; GENERAL INFORMATION:
; APPLICANT: Patel, Salil D
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jenö

; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
; TITLE OF INVENTION: Smooth Muscle Cells
; FILE REFERENCE: 106482.287
; CURRENT APPLICATION NUMBER: US/09/457,646
; CURRENT FILING DATE: 1999-12-09
; EARLIER APPLICATION NUMBER: 60/122,974
; EARLIER FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 60/163,682
; EARLIER FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7

LENGTH: 1143

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-646-7

Alignment Scores:

Pred. No.: 0.0983 Length: 1143
Score: 125.50 Matches: 80
Percent Similarity: 37.01% Conservative: 24
Best Local Similarity: 28.47% Mismatches: 88
Query Match: 9.62% Indels: 90
DB: 4 Gaps: 16

US-09-993-808B-2 (1-256) x US-09-457-646-7 (1-1143)

```
QY 9 ArgGlyAlaAlaGlyAlaGlu-----ValAlaAlaValGluValThrGlnValValGly 26
Db 105 AGAGGAGGTGCGGCGCTGCTGGAGCGGGCGCTGCCAACGCCACCGA-ATAGTTACG 163
QY 27 ValArgThrArgSerArgSer-----AlaAla-AlaThrGlyGlyValAlaLysVa 43
Db 164 GTGGAGGCGATCCAGTTCATGATGATGAGCGCGCGCGAGTGGCGAGCTGCT--GC 220
QY 43 lAlaProArgArg-----Ly 48
Db 221 TGCTCCACGGCGGAGGCCAACCTGCGCGCCCGCCACTCTCACCCGACCGGTGCACG 280
QY 48 sArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlySerCy 68
Db 281 ACGTGTCGCGGAGGGCTTCT-----GGACACGCTGGTGTGCT--GC 320
QY 68 sTyrIleHisLeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerVa 88
Db 321 -----GCACCGGGCGCG----- 332
QY 88 lAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGln----- 104
Db 333 -----GGCGCGCTGGACGTGCGCGCGTGCCT 358
QY 105 -----GlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAl 120
Db 359 GGGGCGCTGCGCGGTGGAGTCCGTCGAGGAGTGGCCATCGCATGTGCGACGGTACC 418
QY 120 aSerSerValAsnLeuGlyGlyGln-ArgGlySerHisThrCysArgSerTyrA 140
Db 419 TGGCGC-----GGCTGCGGGGGGCGCACGAGCGAGTAAACCATGCCCGC--ATAG 466
QY 140 spAlaAlaGluAlaGlyGlyAspHisValLeuValValSerAlaAlaSer---AsnS 159
Db 467 ATCCCGGAGAGTCCCTCAGAC-----ATCCCGATGTGGGGGGGCGAGCGGGGTG 520
QY 159 erGlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyG 179
Db 521 GCGGTTCCGCGGGGTGGATCCGTCAGTCAACCGTGGAGTGTCTAACGGGCGCCCTA 580
```

179 luLeuSerAspLeuCluSerAspLeuAlaGlyHisLysThrClyProSerLeuProAlaA 199
581 GCCTGGAGCGGATGACGCGCAGGCGGAGCAC-----CCCAAGC 622
199 laThrProAlaAlaGluLeuLeuValProProAlaHisGlu-----ileG 214
623 CCTCGCCTGAGGAACTCTTCGCGCGGTGGACCAAGAGTTAACCGGGACTTGG 682
214 lngluPhePheAlaAlaGluAlaGluAlaGlnAlaLysArgPheAlaSerLysTyrAsn 234
683 AGAAGCACTGCAGACATGGAAGAGGCGGAGCGC-----AAGTGGAAAT 730
234 heAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrpAlaProVal 253
731 TCGATTTTCAGATCACAACCCCTAGAG-----GGCAAGTACGAGTGGCAGAGGTG 783

RESULT 14

US-08-902-572-5

; Sequence 5, Application US/08902572

; Patent No. 6495526

; GENERAL INFORMATION:

; APPLICANT: Gyuris, Jeno

; APPLICANT: Lamphere, Lou

; APPLICANT: Beach, David H.

; TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND

; TITLE OF INVENTION: RELATED THERETO

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FOLEY, HORAG & ELIOT LLP

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109-2170

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/902,572

; FILING DATE: 29-JUL-1997

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Vincent, Matthew P.

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: MIV-069.03

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-832-1000

; TELEFAX: 617-832-7000

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1143 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1140

US-08-902-572-5

Alignment Scores:

Pred. No.: 0.0983 Length: 1143

Score: 125.50 Matches: 80

Percent Similarity: 37.01% Conservative: 24

Best Local Similarity: 28.47% Mismatches: 88

Query Match: 9.62% Indels: 90

DB: 4 Gaps: 16

US-09-993-808B-2 (1-256) x US-08-902-572-5 (1-1143)

9 ArgGlyAlaAlaGlyAlaGlu-----ValAlaAlaValGluValThrGlnValValGly 26
105 AGAGGAGTGGCGGCGCTGCTGGAGCGGGCGCTGCCCAACGACCGA-ATACTTACG 163
27 ValArgThrArgSerArgSer-----AlaAla-AlaThrGlyValAlaLysVa 43
164 GTGCGAGCGCCATCCAGGTCTGATGATGGCGAGCGCCGAGTGGCGGAGTCT- --GC 220
43 lAlaProArgArg-----ly 48
221 TGCTCCACGGCGGAGCCCACTCGCGGACCCCGCCACTCTCACCAGCCCGCTGCAGC 280
48 sArgAlaProAlaGlyProAlaAlaAlaValSerAlaGlyLysArgGlyGlySerCy 68
281 ACGTGGCGGAGGCGCTTCT- --GGACACGCTGGTGGTGGT- -- 320
68 sTyrIleHisLeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerVa 88
321 -----GCACCGGCGCG- -- 332
88 lAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGln----- 104
333 -----GGCGGCTGGACGTGCGCGATGCT 358
105 -----GlnGlyAlaAlaAlaGlyLeuSerArgCysSerSerThrAl 120
359 GGGCGCGTCTGCCCGTGGACCTGGTGGAGGAGTGGCGCATCGCATGTGCGACGTACC 418
120 aSerSerValAsnLeuGlyGlyGln-ArgGlySerHisThrCysArgSerTyrA 140
419 TCGCGC-----GGTCTGGGGGCGACCGAGAGGAGTAAACCATGCCCGC- --ATAG 466
140 sAlaAlaGluAlaGlyGlyHisValLeuValAspValSerAlaAlaSer- --AsnS 159
467 ATGCGCGGAGGTCCCTCAGAC-----ATCCCGATGCTGGCGGGGCGACGGGGG 520
159 erGlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyG 179
521 GCGGTTCGCGGGGGTGGATCCGTGAGTCAAAGCTGCGAGTGTCTTAACGGGCGCCTA 580
179 luLeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaA 199
581 GCCTGGAGCGGATGACCGCCAGCGGAGCGGAGCAC-----CCCAAGC 622
199 laThrProAlaAlaGluLeuValProProAlaHisGlu-----ileG 214
623 CCTCGCGCTGCAGAACCTCTTCGCGCGGTGGACCAAGAGTTAACCGGGACTTGG 682
214 lngluPhePheAlaAlaGluAlaGluAlaGlnAlaLysArgPheAlaSerLysTyrAsn 234
683 AGAAGCACTGCAGACATGGAAGAGGCGGAGCGC-----AAGTGGAAAT 730
234 heAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrpAlaProVal 253
731 TCGATTTTCAGATCACAACCCCTAGAG-----GGCAAGTACGAGTGGCAGAGGTG 783

RESULT 15

US-09-516-065-7

; Sequence 7, Application US/09516065

; Patent No. 6521602

; GENERAL INFORMATION:

; APPLICANT: Patel, Salil D

; APPLICANT: McArthur, James G

; APPLICANT: Gyuris, Jeno

; APPLICANT: Mendez, Michal

; APPLICANT: Piner, Mitchel

; TITLE OF INVENTION: Anti-Neoplastic Compositions and Uses Thereof

; FILE REFERENCE: Cell 406; 106482.632

; CURRENT APPLICATION NUMBER: US/09/516,065

; CURRENT FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 60/128,515

; PRIOR FILING DATE: 1999-04-09

; PRIOR APPLICATION NUMBER: 60/128,271

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: October 2, 2004, 11:11:40 ; Search time 447 Seconds
(without alignments)

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVPLDAGGRFFWAPVVS 256

Scoring table:	BL0SUM62
	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Zgapop 6.0 , Zgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

MODEL=frame_82n.model _DEV=xlp
Q=/cgnr1/USFSTC_spool_p/US09938308/runat_01102004_171704_131115/app_query.fasta_1.455
DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blowsum62
TECL=human40.cdi -LIST=45 -DOCALIGN=200 -THR MAX=100
THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
MAXLEN=200000000 -USER=US09938308_@C@N_1_723_@runat_01102004_171704_131115
NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LOGLONG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREATS=1 -XGAPOP=10 -XGAPEXT=0.5
FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELBEXT=7

```

```

Database : Published Applications NA:**
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Length	DB	ID	Description
1	1	100	100	1	1	1
2	2	100	100	2	2	2
3	3	100	100	3	3	3
4	4	100	100	4	4	4
5	5	100	100	5	5	5
6	6	100	100	6	6	6
7	7	100	100	7	7	7
8	8	100	100	8	8	8
9	9	100	100	9	9	9
10	10	100	100	10	10	10
11	11	100	100	11	11	11
12	12	100	100	12	12	12
13	13	100	100	13	13	13
14	14	100	100	14	14	14
15	15	100	100	15	15	15
16	16	100	100	16	16	16
17	17	100	100	17	17	17
18	18	100	100	18	18	18
19	19	100	100	19	19	19
20	20	100	100	20	20	20
21	21	100	100	21	21	21
22	22	100	100	22	22	22
23	23	100	100	23	23	23
24	24	100	100	24	24	24
25	25	100	100	25	25	25
26	26	100	100	26	26	26
27	27	100	100	27	27	27
28	28	100	100	28	28	28
29	29	100	100	29	29	29
30	30	100	100	30	30	30
31	31	100	100	31	31	31
32	32	100	100	32	32	32
33	33	100	100	33	33	33
34	34	100	100	34	34	34
35	35	100	100	35	35	35
36	36	100	100	36	36	36
37	37	100	100	37	37	37
38	38	100	100	38	38	38
39	39	100	100	39	39	39
40	40	100	100	40	40	40
41	41	100	100	41	41	41
42	42	100	100	42	42	42
43	43	100	100	43	43	43
44	44	100	100	44	44	44
45	45	100	100	45	45	45
46	46	100	100	46	46	46
47	47	100	100	47	47	47
48	48	100	100	48	48	48
49	49	100	100	49	49	49
50	50	100	100	50	50	50
51	51	100	100	51	51	51
52	52	100	100	52	52	52
53	53	100	100	53	53	53
54	54	100	100	54	54	54
55	55	100	100	55	55	55
56	56	100	100	56	56	56
57	57	100	100	57	57	57
58	58	100	100	58	58	58
59	59	100	100	59	59	59
60	60	100	100	60	60	60
61	61	100	100	61	61	61
62	62	100	100	62	62	62
63	63	100	100	63		

ALIGNMENTS

RESULT 1

```

US-09-93-308-1
; Sequence 1, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

```

NAME/KEY: CDS
LOCATION: (134)...(904)
OTHER INFORMATION:
US-09-993-308-1

Alignment Scores:
Pred. No.: 2,2e-118 Length: 1372
Score: 1304.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-993-308-1 (1-1372)

```
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 134 ATGGGGAAGTACATGCGCAAGTGCAGGGGCGCGCAGCGGAGGTCGCGCGCGTCGAG 193
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyVal 40
Db 194 GTTACGAGGTCGTCGCGCTCCGAGCGAGGTCCAGGTCGCGGCGCGGACCGCGGTGTC 253
QY 41 AlaLysValAlaProArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
Db 254 GCGAAGGTGCGCCCGAGGAGGAGGCGCGCGCGCGGAGGCTGCTGCGCGCGTGC 313
QY 61 AlaGlyLysAspGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
Db 314 GCTGTGGGACCGCGGAAGTGTACATCCACTCGGTAGCGGAGCCACCTCGCTCTACGAC 373
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAspGly 100
Db 374 CCGCTCAGCGCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGlyLeuSerArgCysSerThrAla 120
Db 434 GCTGAGGACAGACGAGGCGCGGCTCGCGCGCGGCTCGCGCGGAGGCTGCTGCTGCTGCT 493
QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
Db 494 TCGTCTGAGTCTGCGGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
QY 141 AlaAlaGluAlaGlyCysAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
Db 554 GCTGAGGAGCTGCGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
QY 161 SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
Db 614 AGCGGCCACAGCCGAGAGGCGAGAGACGACCCATCGAGCGCGGCGCGGAGCTC 673
QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
Db 674 AGCGATCTGAGTCTGCGGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
QY 201 ProAlaAlaGluLeuLeuValProProAlaHisGluLeuGlnPhePheAlaAlaAla 220
Db 734 CCGGCTGCGGAGTCTGCGGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
Db 794 GAGGCGGCCAGGCCAAGCGCTTGTCTCCAGTACACTTCGACTCGCTGCGCGCGGTG 853
QY 241 ProLeuAspAlaGlyArgPheGluThrAlaProValValSerIle 256
Db 854 CCGCTCAGCGCGCGCGGTCGAGTGGCGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
```

RESULT 2

US-09-993-808B-1

Sequence 1, Application US/09993808B
Publication No. US20040003433A1
GENERAL INFORMATION:
APPLICANT: Gordon-Kamm, William

APPLICANT: Lowe, Keith
APPLICANT: Sun, Yuejin
APPLICANT: Dilkes, Brian
APPLICANT: Larkins, Brian
TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
TITLE OF INVENTION: and Uses Thereof
FILE REFERENCE: 1146
CURRENT APPLICATION NUMBER: US/09/993,808B
CURRENT FILING DATE: 2001-11-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1372
TYPE: DNA
ORGANISM: zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (134)...(902)
US-09-993-808B-1

Alignment Scores:

Pred. No.: 2,2e-118 Length: 1372
Score: 1304.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-993-808B-1 (1-1372)

```
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 134 ATGGGGAAGTACATGCGCAAGTGCAGGGGCGCGCAGCGGAGGTCGCGCGCGTCGAG 193
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyVal 40
Db 194 GTTACGAGGTCGTCGCGCTCCGAGCGAGGTCCAGGTCGCGGCGCGGACCGCGGTGTC 253
QY 41 AlaLysValAlaProArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
Db 254 GCGAAGGTGCGCCCGAGGAGGAGGCGCGCGCGGAGGCTGCTGCGCGCGTGC 313
QY 61 AlaGlyLysAspGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
Db 314 GCTGTGGGACCGCGGAAGTGTACATCCACTCGGTAGCGGAGCTGCTGCTGCTGCTGCT 373
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
Db 374 CCGCTCAGCGCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGlyLeuSerArgCysSerThrAla 120
Db 434 GCTGAGGACAGACGAGGCGCGGCTCGCGCGCGGCTCGCGCGGAGGCTGCTGCTGCTGCT 493
QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
Db 494 TCGTCTGAGTCTGCGGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
QY 141 AlaAlaGluAlaGlyCysAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
Db 554 GCTGAGGAGCTGCGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
QY 161 SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
Db 614 AGCGGCCACAGCCGAGAGGCGAGAGACGACCCATCGAGCGCGGCGCGGAGCTC 673
QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
Db 674 AGCGATCTGAGTCTGCGGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
QY 201 ProAlaAlaGluLeuLeuValProProAlaHisGluLeuGlnPhePheAlaAlaAla 220
Db 734 CCGGCTGCGGAGTCTGCGGCGGAGTCCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
QY 201 ProAlaAlaGluLeuLeuValProProAlaHisGluLeuGlnPhePheAlaAlaAla 220
Db 734 CCGGCTGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
```

QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
 Db 794 GAGCGGCCCAAGCCAGCGCTTTCCTCCAAAGTACAACTTCGACTTCGTCGGCGGGGTG 853

QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
 Db 854 CCCCTCGACGCGCGCGCGGTTCGAGTGGCGCGCGGTGGTCAGCATC 901

RESULT 3

US-10-425-114-34947
 ; Sequence 34947, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 2003-04-28
 ; SEQ ID NO 34947
 ; LENGTH: 1111
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLMO17355P05_FLI
 US-10-425-114-34947

Alignment Scores:
 Pred. No.: 5,18e-110 Length: 1111
 Score: 1217.50 Matches: 245
 Percent Similarity: 96.50% Conservative: 3
 Best Local Similarity: 95.33% Mismatches: 4
 Query Match: 93.37% Indels: 5
 DB: 13 Gaps: 3

US-09-993-808B-2 (1-256) x US-10-425-114-34947 (1-1111)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
 Db 32 ATGGGAAGTACATGCGCAAGTCAGGGGGCGCGCAGCGCGGAGGTGCGCGCGTCGAG 91

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
 Db 92 GTTACGACGTCGTCGCGCTCCGACAGGTCCAGGTCCGGCGCGCGACCGCGGTGTC 151

QY 41 AlaLys---ValAlaProArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
 Db 152 GCGAAGTCTCGCCCGCAGGAGAGAGCGCGCGCGCGCGCGCTGCTGCTCGCTG 211

QY 60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
 Db 212 GCGCGTGTGGGACGCGGAGCTGCTATCACTGCGGTAGCCGCATGCTGTTTCATG 271

QY 80 AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
 Db 272 GCACCGCTCAGCGCAGCGC-----CCGTGCTCCGACCCCGCGGAGGTGCTGAT 325

QY 100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerThr 119
 Db 326 GGGCTGTCAGGACAGCGCGCGCTCGCGCGCGCTCGCGCGCTGCTGCTCCAGCAGC 385

QY 120 AlaSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyr 139
 Db 386 GCGTCGTGGTGAC-----GTGGGGGTGACGCGGGAGCCACACTCCCTCCGAC 439

QY 140 AspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSer 159

Db 440 GACCTCGACGAGGTCGGCGGATCATCGTCTGTGTGNTGTCTCGCGCGCGCAACTCC 499
 QY 160 GlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGlu 179
 Db 500 GGGAGCGGCCAGACCCGAGAGCGGAGAGACGACCATCGAGCCGGCGCGCACGGCGAG 559
 QY 180 LeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAla 199
 Db 560 CTCAGCGATCTGGAGTCGATCTGCGCGGGGCACAAGACTGCGCGCTCGCTACCGCGCGCA 619
 QY 200 ThrProAlaAlaGluLeuValProProAlaHisGluIleGlnGluPhePheAlaAla 219
 Db 620 ACGCGCGCTCGGAGCTGATCGTCCGCCAGACACAGATCCAGGAGTCTTCGCCGCC 679
 QY 220 AlaGluAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGly 239
 Db 680 GCCGAGCGGCCAGGCCAAGCGCTTTCCTCAAGTACAACCTTCGACTTCGTCCGTGGC 739
 QY 240 ValProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
 Db 740 GTGCGCTCGACGCGCGCGCGGTTCGAGTGGCGCGCGGTGCTCAGCATC 790

RESULT 4

US-10-425-114-18379/c
 ; Sequence 18379, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 2003-04-28
 ; SEQ ID NO 18379
 ; LENGTH: 985
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI
 US-10-425-114-18379

Alignment Scores:
 Pred. No.: 1,34e-94 Length: 985
 Score: 1059.50 Matches: 213
 Percent Similarity: 96.00% Conservative: 3
 Best Local Similarity: 94.67% Mismatches: 4
 Query Match: 81.25% Indels: 5
 DB: 13 Gaps: 3

US-09-993-808B-2 (1-256) x US-10-425-114-18379 (1-985)

QY 33 SerAlaAlaAlaThrGlyGlyValAlaAlaLys---ValAlaProArgArgLysArgAlaPro 51
 Db 984 TCCGCGCGCGCAGCGCGGTGTCGGAAGTGTGCGCCCGGAGGAGAGAGGGCGCG 925

QY 52 AlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHis 71
 Db 924 GCGGGGAGCTCTGCTCGCTCGTGGCGCTGTGTGGGACGCGCGAAGCTGCTATCATCC 865

QY 72 LeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerVal 91
 Db 864 CTGCTAGCGCATGCTGTTCATGGCAGCCGCTCAGCCGAGCCG-----CCGTGGTT 811

QY 92 ProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAla 111
 Db 810 CCGACCCCGCGGAGGCTGCTGATGCGCTGACAGACGACGAGGCGCGGCTCGCGGCC 751

Qy	112	GlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlnArg	131
Db	750	GGGCTCTCGCGCTTGTCTCAGACACGCGTCTCGGTGCAC-----GTGGGGGGTCAGCGC	697
Qy	132	GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuVal	151
Db	696	GGGAGCCACACCTCGCGCTTCGACGACGCTGCAGAGGCTGGCGGGGATCAGTCTCTGGTG	637
Qy	152	AspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThr	171
Db	636	GATGTCCTCGCGCGCGAGCAACTCTCGGGAGCGGCCCCAGACCGCGAGAGGCGAGACGACG	577
Qy	172	ProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys	191
Db	576	CCATCGAGCCGGGGCCACGGCGAGCTCAGCGATCTGGAGTCGATCTGGCGGGGCACAAAG	517
Qy	192	ThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuValProProLahis	211
Db	516	ACTGGCCGCTCGCTTACCGGGCGCAACGCCGGCTGCGGAGCTGATCGTCGCGCCACACAC	457
Qy	212	GluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLys	231
Db	456	GAGATCCAGGAGTTCTTCGCGCGCCCGAGCGGCCACAGGCGCAAGCGCTTTGCTTCCAAAG	397
Qy	232	TyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrpAla	251
Db	396	TACAACTTCGACTTCGTCGTGGCTGCCCTTCGACGCGCGCGCGCTTCGAGTGGCGG	337
Qy	252	ProValValSerIle	256
Db	336	CCGGTGGTTCAGCATC	322

RESULT 5

```

US-10-333-006-9
; Sequence 9, Application US/10333006
; Publication NO. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Drouail, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz,, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; CURRENT APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-333-006-9

```

Alignment Scores:	
Pred. No.:	2,958-56
Score:	669.00
Percent Similarity:	67.4%
Best Local Similarity:	60.30%
Query Match:	51.30%
DB:	16
	Gaps: 10
	Indels: 20
	Mismatches: 67
	Conservative: 19
	Matches: 161
	Length: 1242

US-09-993-808B-2 (1-256) x US-10-333-006-9 (1-1242)

1 MetGlvLvsTyrMetArqLysCysArqGlyAlaAlaGluValAlaAlaValGlu 20

Db	18	ATCGGCAAGTACATGAGAAATTTCAGGGGGCCACGGGGGAGGAGTTGGCCGCATGGAG	77
Qy	21	ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly---	39
Db	78	GTCAAGCAGGTGTTGGCGTCCGACAGAGTCCAGGTCGGCAGCGCGCGCGCCGACG	137
Qy	40	-----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu	54
Db	138	ACGACGAAGGTGCAGGGCGCGTCGGCGCGTCCACCAGGAGGAGGAAGCGCTGCTCGCG	197
Qy	55	ProAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSer	74
Db	198	ACGCGCGTCTGGGGACTACTCGCGCTGACGCGCGGAGCTGCTACTCCAGCTGAGGAGC	257
Qy	75	ArgMetLeuPheMetAlaProGlnProGlnProSerValArgSerValProThrPro	94
Db	258	CGCATGCTGTTTCATGGCCCCCGGAGGCGCGCGCGCGCG---AGGGCTCCGGTTGTA	314
Qy	95	ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSer	114
Db	315	GCGGAGCGCGG---GGTCCGGGAACGGAGCGCGCGCATCGCGCGCTGGCCTCTCG	371
Qy	115	ArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHis	134
Db	372	CGTTGCTCCAGCACCGCGCTGCTCGTGGAC-----GCGGGCGGCTCAGCACAGGAGCCCTC	425
Qy	135	ThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSer	154
Db	426	GCGTGCCCGCTCC---GACGTCCGGAGCGACGACGAGCAT-----GTCCCGGAGGCG	476
Qy	155	AlaAlaSerAsnSerGlySerGlyProAspArgGluArgArgGluThrThrProSerSer	174
Db	477	TCCCGAGAGCACTCGCGAGCGCGCGGTGACCGCGAGAGGAGAGAACAACTCCATCAAGC	536
Qy	175	ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAla---GlyHisLysThrGly	193
Db	537	TTTCTCCCCGCGAGGTGAGCGAICTGGAGTCGGATCTCGGTGGAGGACAGAAGCGCAGC	596
Qy	194	ProSerLeuPro---AlaAlaThrProAlaAlaGluLeuLeu-----ValPro	208
Db	597	CGTCCACTACCTCTCGGGCAACAGCTCCAGCACAGCAAGCCACCGCGCCGAGAGATCCG	656
Qy	209	ProAlaHisGluIleGlnPhePheAlaAlaGluAlaAlaGlnAlaLysArgPhe	228
Db	657	CCGCGCCCGCAGATCGAGGCGCTTCTTCGCGCGCGCGGAGGAGCTGAGCCCAAGCGCTTC	716
Qy	229	AlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPhe	248
Db	717	GCGCCCAAGTACAACTTCGAGCTCGTTCGCGCGCTGCCCTCGACGCC---GGTCGGTTC	773
Qy	249	GluTrpAlaProValValSer	255
Db	774	GAGTGGACTCCGGTGGTCAGC	794

RESULT 6

```

US/10-437-963-92632/c
Sequence 92632, Application US/10437963
Publication No. US20040123342A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yingwei
APPLICANT: Cao, Yonghui
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules
FILE REFERENCE: Plants and Uses Thereof
CURRENT APPLICATION NUMBER: US/10-437,963
CURRENT FILING DATE: 2003-05-14

```



```
/ NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 92632
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1
US-10-437-963-92632

Alignment Scores:
Pred. No.: 4.04e-56 Length: 1671
Score: 669.00 Matches: 161
Percent Similarity: 67.42% Conservative: 19
Best Local Similarity: 60.30% Mismatches: 67
Query Match: 51.30% Indels: 20
DB: 17 Gaps: 10

US-09-993-808B-2 (1-256) x US-10-437-963-92632 (1-1671)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaGluValAlaAlaValGlu 20
DB 1206 ATGGGCAAGTACATGAGGAATTCAGGGGGGCCACGGGGGAGGATTGGCCGCATGGAG 1147
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly--- 39
DB 1146 GTCACGCGGTGTTGGCGTCCGACGAGTCCGAGGTCGAGGTCGCGCGCGCGCGCGGCGGCG 1087
QY 40 -----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu 54
DB 1086 ACGACGAAGGTGAAGCGCGTCCGCGCGCTCCACGAGGAGGAGGAGGCGCTGCTGCCG 1027
QY 55 ProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSer 74
DB 1026 ACGCGGTGCTGGGGACTACTCGCGGTGACGGCGGGAGCTGCTACCTCCAGCTGAGGAGC 967
QY 75 ArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValProThrPro 94
DB 966 CGCATGCTGTTCATGGCCCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 910
QY 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlnGlnAlaAlaAlaGlyLeuSer 114
DB 909 GCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 853
QY 115 ArgCysSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGlySerHis 134
DB 852 CGTGTCTCAGACGCGCTGCTCGGTGAC-----CGCGCGCTCAGGACAGGAGCCTC 799
QY 135 ThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
DB 798 GCGTGCCTGCCCTCC---GACGTCGCGGAGGCGAGGCGAGCGAT-----GTCCCGGAGGCG 748
QY 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGlnThrThrProSerSer 174
DB 747 TCGCGAGGACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 698
QY 175 ArgAlaHisGlyGluLeuSerAspLeuLeuSerAspLeuAla---GlyHisLysThrGly 193
DB 687 TTTCTCCCGGAGAGTGAGCGATCTGGAGTCGATCTGCTGGAGGACAGAGCGGAGC 628
QY 194 ProSerLeuPro---AlaAlaThrProAlaAlaGluLeuIle-----ValPro 208
DB 627 CGTCCACTACTTCTGCGGCAACAGCGCTCAGCAGCAAGCCAGCGCGCGCGCGCGCGCGCGCG 568
QY 209 ProAlaHisGluIleGlnGlnPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPhe 228
DB 567 CCGCGCGCGCGAGATCGAGGCGTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 508
QY 229 AlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPhe 248
DB 507 GCGCGCAAGTACAACTTCGAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 451
QY 249 GluTrpAlaProValValSer 255
|||||
```

```
Db 450 GAGTGGACTCCGGTGGTCAGC 430
RESULT 7
US-10-425-114-18085/c
; Sequence 18085, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18085
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3078-022-G5_FLI
US-10-425-114-18085

Alignment Scores:
Pred. No.: 2.16e-56 Length: 834
Score: 668.50 Matches: 174
Percent Similarity: 76.05% Conservative: 7
Best Local Similarity: 73.11% Mismatches: 30
Query Match: 51.27% Indels: 27
DB: 13 Gaps: 17

US-09-993-808B-2 (1-256) x US-10-425-114-18085 (1-834)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaGluValAlaAlaValGlu 20
DB 696 ATGGGAAGTACATGCGCAAGCGCGGGGGCGCGCGGGGGTGGCGCGAGTCAG 637
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThr---GlyGly 39
DB 636 GTCTCGCAGGTCTCGCGGTCCGAGGAGGTCCAGGTCGCGCGCGCGCGCGCGCGCGGT 577
QY 40 ValAlaLysValAla---ProArgArgLysArgAla-----ProAla----- 52
DB 576 GTCGCGAAGTCTCGTCCGCGAGGAGGAGGAGGCGCTGCTGCCCGCGCGAGCAGC 517
QY 53 -----GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySer---CysTyr 69
DB 516 GCGTCGGGGAGCGCTGCTGCCGCGGTGGGGTGGGGAGCGCGGAGCTGCTGTCTAC 457
QY 70 IleHisLeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAsp 89
DB 456 ATCCACTCGGAGCGCATGCTGTTTCATGCGACGACCTCAGCAGCAACCGTCG----- 403
QY 90 SerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGly---AlaAla 108
DB 402 GCGGCTCTGACGCGCGTGGAGGCTGCT---GGTGCAGCACAGCAGGCGCGGTGGTGGC 346
QY 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGly 128
DB 345 CTCGCGGCTGGCCTCTCGCGTCTCAGCACGCGCTGCTCGCTCCGCTCCGCTGCGGAGT 292
QY 129 GlyGlnArg-----GlySerHisThrCysArgSerTyrAspAla---AlaGluAla 144
DB 291 GGCCAGCAGCCCGGAGGAGGAGCGCGCTGCTGCGCTCCGCTCCGCTGCGGCTGGAGT 232
QY 145 GlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAsp 164
DB 231 GACGGGATCACGTC---CCGATGTGTCACCGCGAGCAACTCGGGGAGCGTCCCGGAC 175
```

QY 165 ArgGluArgGluThrPro---SerSerArgAlaHis---GlyGluLeuSerAsp 182
Db 174 CGCGAGAGAGAGAGACGACGACATCGTCGAGCGGGCGCACGGCGGAGCTCAGCGAT 115
QY 183 LeuGluSerAspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---Ala 199
Db 114 CTGAGTCTGGATCTGCTGGGGCGGCGAGAGACTGGCTGCTGCTCGCGCGCGCAACA 55
QY 200 ThrProAlaAlaGluLeuValProProAlaHisGluLeuGlnGluPhePhe 217
Db 54 ACCTGGCTGGAGAGCTATCTGTCGCCGACGACAGGAGATCCAGGAATCTTTC 1

RESULT 8

US-09-993-308-5
; Sequence 5, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent version 3.1
; SEQ ID NO 5
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(839)
; OTHER INFORMATION: The 'r' at location 491 stands for g or a.
US-09-993-308-5

Alignment Scores:

Pred. No.: 6e-56 Length: 841
Score: 664.00 Matches: 172
Percent Similarity: 75.11% Conservative: 6
Best Local Similarity: 72.57% Mismatches: 31
Query Match: 50.92% Indels: 28
DB: 9 Gaps: 16

US-09-993-808B-2 (1-256) x US-09-993-308-5 (1-841)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 159 ATGGGAAGTACATCGCAGCGAGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGG 218
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThr---GlyGly 39
Db 219 GTCTCGAGGTCTGTCGGGTCGCGAGAGTCCAGTCCGCGCGCGCGCGCGCGCGGT 278
QY 40 ValAlaLysValAla---ProArgLysArgAla-----ProAla----- 52
Db 279 GTCCGAAGTCTGCTCCGCCGAGGAGAGAGAGGCGTCTGCCCGCGCGAGCGTGGAGC 338
QY 53 -----GlyGluProAlaAlaValSerAlaGlyGlyAspGlyGlySer---CysTyr 69
Db 339 ACCTCGGGGACCTGCTGCTGGCGCTGGTGGTGGGAGCGCGGAGCTGCTGCTAC 398
QY 70 IleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAsp 89
Db 399 ATCCACCTCGGAGCGCATGCTGTTTCATGGCAGCACCTTCAGCAGCAACCGTCG----- 452
QY 90 SerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGly---AlaAla 108
Db 453 GCGGCTCTACCCCGGTGGAGCTGCT---GGTGGCGCACACACAGCGGGGTGGGCG 509

QY 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGly 128
Db 510 CTGCGGGCTGGCTCTCGGTTGCTCCAGCAGCGCGTCTGCTGGTGGACGTC----- 560
QY 129 GlyGlnArgGlySerHisThrCysArgSerTyrAspAla---AlaGluAlaGlyGlyAsp 147
Db 561 -----GGGGGCGCAGCGCTGCCGCTCCGACGCTGCGCGCTGCGAGGTTTACCGGGAT 611
QY 148 HisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArg 167
Db 612 CACGTC-----CCGATGTCGTACCCGAGCAACTCGGGAGCGTCCCGACCGCGAGAG 668
QY 168 ArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeuGluSer 185
Db 669 AGAGAGACGACGCGCATCTGTCGAGCGGGCGCACGCGCGGAGCTCAGCGATCTGGAGTCG 728
QY 186 AspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThrProAla 202
Db 729 GATCTGGTGGGGCGGAGAGACTGGCTGCTGCTCGCGCGCGACACACATCGGCT 788
QY 203 AlaGluLeuValProProAlaHisGluLeuGlnGluPhePheAlaAla 219
Db 789 GCGGAGCTGATCTGTCGCCGACGACGAGATCCAGGAATCTTCGCGGCC 839

RESULT 9

US-09-993-808B-5
; Sequence 5, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(841)
; OTHER INFORMATION: r= g or a
US-09-993-808B-5

Alignment Scores:

Pred. No.: 6e-56 Length: 841
Score: 664.00 Matches: 172
Percent Similarity: 75.11% Conservative: 6
Best Local Similarity: 72.57% Mismatches: 31
Query Match: 50.92% Indels: 28
DB: 11 Gaps: 16

US-09-993-808B-2 (1-256) x US-09-993-808B-5 (1-841)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 159 ATGGGAAGTACATCGCAGCGAGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGG 218
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThr---GlyGly 39
Db 219 GTCTCGAGGTCTGTCGGGTCGCGAGTCCAGTCCGCGCGCGCGCGCGCGGT 278
QY 40 ValAlaLysValAla---ProArgLysArgAla-----ProAla----- 52
Db 279 GTCCGAAGTCTGCTCCGCCGAGGAGAGAGAGGCGTCTGCCCGCGCGAGCGTGGAGC 338

QY 53 -----GlyGluProAlaAlaValSerAlaGlyGlyAspGlyGlySer---CysTyr 69
Db 339 ACCTCGGGGAGCTGTGTGGCGGGCGTGGTGGTGGGACGGCGGAGAGCTGTGTGTAC 398
QY 70 IleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAsp 89
Db 399 ATCCACCTCGGAGCGCATGCTGTTCATGGCAGCACCTCAGCAGCAACCGTGC-----452
QY 90 SerValProThrProValGluAlaAlaAspGlyAlaGlyGlnGlnGly---AlaAla 108
Db 453 CGCGCTCTGACGGCGGTGGAGGCTGCT---GGTGGCGCACARCAAGCGGGGTGGTGGCG 509
QY 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGly 128
Db 510 CTGGCGCTGGCTCTCGCTGCTCCAGCAGCGCGTCTGGTGGAGCTC-----560
QY 129 GlyGlnArgGlySerHisThrCysArgSerTyrAspAla---AlaGluAlaGlyGlyAsp 147
Db 561 -----GGGGCCACCGCTGCGCTCCGACGCTCGCGTGGAGGTGACGGGAT 611
QY 148 HisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArg 167
Db 612 CACGTC---CCGGATGTCTCACCGCAGCAACTCCGGGAGCGTCCCGGACCGCGAGAGG 668
QY 168 ArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeuGluSer 185
Db 669 AGAGACGACGACCATCTGTGAGCGGGCGCACGGCGGAGCTCAGCATCTGGAGTCG 728
QY 186 AspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThrProAla 202
Db 729 GATCTGGTGGCGGAGAGACTGGCTGCTGCTGCTGCGCGCGGACCAACACATCGGCT 788
QY 203 AlaGluLeuLeuValProProAlaHisGluLeuGlnGlnPhePheAlaAla 219
Db 789 GCGAGCTGATCTGCGCGCCAGCACAGGAGATCCAGGATTTCTTCGCGGCC 839
RESULT 10
US-10-767-701-24179
; Sequence 24179, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24179
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30165606
US-10-767-701-24179
Alignment Scores:
Pred. No.: 2,69e-41 Length: 510
Score: 512.00 Matches: 112
Percent Similarity: 85.82% Conservative: 3
Best Local Similarity: 83.58% Mismatches: 5
Query Match: 39.26% Indels: 14
DB: 17 Gaps: 5
US-09-993-808B-2 (1-256) x US-10-767-701-24179 (1-510)
QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaVal 19
Db 121 ATGGGGAAGTACATGCGCAGCAAGTGCAGGGGGCGCGGAGAGGTGCGCGCAGTC 180

QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db 181 GAGGTACGCGAGGTGTTCGGCGTTCAGACGAGGTCCAGGTCCGGCGGCGGACCGGGGT 240
QY 40 ValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 241 GTCGGAAGGTCTGCTCGAGGAGGAAGAGGCGGTGACGCGCGCGGAAACGTGCCGCG 300
QY 53 GlyGluProAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 301 GGGGAGCT---GCCGCGTGGCGTGGCGGAGCGCGGAGCTCTACATCCACTG 357
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 358 CGAAGCGCATGCTGTTCATGGCACCGCTCAGCCGCGAGCGG-----TCGGTTCG 408
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
Db 409 ACGCCCGCGAGGCTGCT-----GGTGACGACAGCAGGCGCGGCTCGTGGCTGG 462
QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGly 126
Db 463 CTCTCGGCTTGTCTCCAGCACGCGTCTCGTGGAGCTCGG 504
RESULT 11
US-10-333-006-5
; Sequence 5, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n = A, T, C or G
US-10-333-006-5
Alignment Scores:
Pred. No.: 3.12e-30 Length: 568
Score: 399.50 Matches: 95
Percent Similarity: 81.90% Conservative: 0
Best Local Similarity: 81.90% Mismatches: 14
Query Match: 30.64% Indels: 7
DB: 16 Gaps: 7
US-09-993-808B-2 (1-256) x US-10-333-006-5 (1-568)
QY 146 GlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArg 165
Db 8 GGGGATCACGTC---CCGGATGTCTGTCNCGCGAGCAACTCGGGAGCGTCCCGGACCG 64
QY 166 GluArgArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeu 183

Db 65 GAGAGGAGAGAGACGCCATCTGTCGAGCGCGGCACGCGCGAGCTCAGCGATCTG 124
 QY 184 GluSerAspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThr 200
 Db 125 GAGTCGGATCTGTCGGGCGGCGAGAGACTGCTGCTGCTGCTGCCCGCGACACACA 184
 QY 201 ProAlaAlaGluLeuValProProAlaHisGluLeuGlnGluPhePheAlaAla 220
 Db 185 TCGGCTGCGGAGCTGTCGTCGCGCAGCAGCAGAGATCCAGGAATCTTCGCGCGCC 244
 QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
 Db 245 GAGCGCGCCCATGCGCAACGCTTTCTTCCAGTACAACTTCGACTTCGTCGCGCGCGTG 304
 QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
 Db 305 CCCCTCGAGCC---GGCGGTTCGATGGACGCCAGGGGTTCAGCATC 349

RESULT 12

US-10-437-963-19949
 ; Sequence 19949, Application US/10437963
 ; Publication No. US20040123343A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 19949

; LENGTH: 1065

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_2535C.1

US-10-437-963-19949

Alignment Scores:

Pred. No.: 1,02e-21 Length: 1065
 Score: 315.50 Matches: 105
 Percent Similarity: 46.62% Conservative: 26
 Best Local Similarity: 37.37% Mismatches: 81
 Query Match: 24.13% Indels: 69
 DB: 17 Gaps: 13

US-09-993-808B-2 (1-256) x US-10-437-963-19949 (1-1065)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
 Db 1 ATGGGGAAG---AAGAAGAAGCGCGAGCGCGCGCGCGAGGAGCGCGGGTGGTG 57
 QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
 Db 58 GTCGGC-----GGCGTCGTCGTCGCGG----- 78
 QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
 Db 79 GCCGCGGTCCCGCAGGAGGGTGTGGCCAGCGCGAGGAGGGTGTGTTGTGGGC 138
 QY 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
 Db 139 CTGGCGGTGGCGGTGGCAGTGGCGAGACGATGGCGAGCGCGGCGGATGCTATCTGCGTCTG 198
 QY 73 ArgSerArgMetLeuPheMetAlaProGlnProGlnProGlnProSerValAspSerValPro 92

Db 199 CGAGCAGGAGGCGCTG----- 213
 QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
 Db 214 ---CCCTTCGTGGCGCGCGCGGTGCTGCTGCGGAGGAGGAGGCGCTCGGTGATTCTG 270
 QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132
 Db 271 GTGGCGGAGGCGCTCTGTCGTCGTCGCGCGCGGTGGAATTTGTTGGCTGTTCTGCT 330
 QY 133 SerHisThrCysArgSerTyrAsp---AlaAlaGluAlaGlyGlyAspHisValLeuVal 151
 Db 331 GAGGAGGAGGCTATGCGCGGAGAGGTTTTCACGAGGAGGCGGAGGATCAC-----GAC 384
 QY 152 AspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThr 171
 Db 385 GAGGAGAGCTCGTTCGGGACTTCGGTTCGGC-----CGGAGAGGAGCGCGAGCAG 438
 QY 172 ProSerSerArg-----AlaHisGlyGluLeuSerAspLeuGluSerAsp----- 186
 Db 439 CGTTCGAGCGCGCGCGCGCGGAGACGCGGACTCGAGCGAGCGCGAGTCAAAACGAGGAG 498
 QY 187 -----LeuAlaGlyHisLysThrGly 193
 Db 499 GCCAAGCAGCAAAATGTCGCGGAGTTCGACGAGCTCAGCAGCTCAGCAGTTCATTTCC 558
 QY 194 ProSerLeuProAlaAlaThrProAlaAlaGluLeuValProPro-----AlaHis 211
 Db 559 -----GCGACGAGGAGGAGCTTCAGGATGATGGCACCGCGCGCGCGGCA 606
 QY 212 GluLeuGlnGluPhePheAlaAlaGluAlaGlnAlaLysArgPheAlaSerLys 231
 Db 607 GAGATCGAGGAGTTCCTCGCGCTCGGAGAGGTCCGAGGCGCGAGCGTTCGCGCGCAAG 666
 QY 232 TyrAsnPheAspPheValArgGlyValProLeuAspAlaGly-----GlyArgPheGlu 249
 Db 667 TACAACCTTCGAGCTGTGTCGCGGCGTGCCTCGACGCGCGCGCGCGCGGCGGTTTCGAA 726
 QY 250 Trp 250
 Db 727 TGG 729

RESULT 13

US-10-437-963-19960/c
 ; Sequence 19960, Application US/10437963
 ; Publication No. US20040123343A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 19960

; LENGTH: 1398

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_2536C.1

US-10-437-963-19960

Alignment Scores:

Pred. No.: 1,91e-21 Length: 1398
 Score: 314.00 Matches: 107
 Percent Similarity: 45.58% Conservative: 27

Best Local Similarity: 36.39% Mismatches: 78
Query Match: 24.08% Indels: 82
DB: 17 Gaps: 14

US-09-993-808B-2 (1-256) x US-10-437-963-19960 (1-1398)

```
Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 1272 ATGGGGAAG---AAGAAGAAGCGCGCGCGCGGAGGAGGCGCGGGTGGT 1216
Qy 21 ValThrGlnValValGlyValArgThrArgSerAlaAlaThrGlyGlyVal 40
Db 1215 GTCCGC-----GGCGTCGTACGCGG-----1195
Qy 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
Db 1194 GCCCGCGTACGCGGAGAGGTGTGCGGAGCGGAGGAGGTGTGGTTGGTGGGC 1135
Qy 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
Db 1134 CGTGGCGGTGCGGTGGCAGTGGCGGAGACGATGGCGGAGGCGGATGCTATCTCGTCTG 1075
Qy 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 1074 CGGAGCAGGAGGCTG-----1060
Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlyAlaAlaLeuAlaAlaGly 112
Db 1059 ---CCCTTCGTGGCGCGCGGTGTGTCGTCGCGGAGGAGGCGCTCGTGATTCG 1003
Qy 113 LeuSerArgCysSerSerThrAlaSerSer-----ValAsnLeuGlyLeuGlyGly 129
Db 1002 GTGCGGAGCGCGTCTGCTGCTGCTGCGCGCGGTGGTAATTTGGGTCTGTTGGT 943
Qy 130 GlnArgGlySerHisThrCysArgSerTyrAspAlaAla-----142
Db 942 GAGGAGGAGGCTATGGCGGAGAGGTGATTCATGAGCGCGCTTGGGATGTGGTGTGTTG 883
Qy 143 -----GluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsn 158
Db 882 CAGTTTTCGCGCAGGAGGCGGAGATCAC-----GACGAGGAGAGCTCCGTCGCGCAC 829
Qy 159 SerGlySerGlyProAspArgGluArgGluThrThrProSerSerArg-----175
Db 828 TCCGCTCGCGC-----CGCGAGAGGCGCGAGCGCGCTCGAGCGCGCGCGCGCG 775
Qy 176 AlaHisGlyGluLeuSerAspLeuGluSerAsp-----186
Db 774 GGAGACGCGGACTCGAGCGAGCGCGGAGTCAAAACGAGGAGGCCAAGCAGCAAAATGTGCGCG 715
Qy 187 -----LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
Db 714 CGAGTTCGACGACCTCAGCAGCTGCATTTTCACCGCGGA-----GCACGACG 667
Qy 201 ProAlaAlaGluLeuIleValProPro-----AlaHisGluIleGluPhePheAla 218
Db 666 AGGAGCTTCAGGATGATGTCACCGCGCGCGCGCGCGGAGATCGAGAGTTCCTCGCC 607
Qy 219 AlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArg 238
Db 606 GCTCGGAGAGGTCGAGGCGGAGCGGTTCGCCCAAGTACAACTTCGACGCTGCTGCGCGC 547
Qy 239 GlyValProLeuAspAlaGly-----GlyArgPheGluTrp 250
Db 546 GCGCTGCGCTCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 505
```

RESULT 14

US-10-333-006-54
; Sequence 54, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius

; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzefeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 54
; LENGTH: 681

; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE: Artificial sequence
; OTHER INFORMATION: probe or primer
; NAME/KEY: misc feature
; LOCATION: (355)..(356)
; OTHER INFORMATION: n = A, T, C or G

US-10-333-006-54

Alignment Scores:

Pred. No.: 2,38e-18 Length: 681

Score: 279.00 Matches: 98

Percent Similarity: 43.51% Conservative: 26

Best Local Similarity: 34.39% Mismatches: 71

Query Match: 21.40% Indels: 90

DB: 16 Gaps: 13

US-09-993-808B-2 (1-256) x US-10-333-006-54 (1-681)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20

Db 1 ATGGGGAAG---AAGAAGAAGCGCGCGCGCGGAGGAGGCGCGGGTGGT 57

Qy 21 ValThrGlnValValGlyValArgThrArgSerAlaAlaThrGlyGlyVal 40

Db 58 GTCCGC-----GGCGTCCGTACGCGG-----78

Qy 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60

Db 79 GCCGCCCTCACGCGGAGGAGGCGTGTGCGGAGCGCGGAGGAGGCGTGTGTTGGTGGC 138

Qy 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72

Db 139 CGTGGCGGTGCGCGTGTGCGTGTGCGGAGCAGTGTGCGGAGGAGGCGGATGCTATCTCGTCTG 198

Qy 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92

Db 199 CGGAGCAGGAGGCTG-----213

Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlyAlaAlaLeuAlaAlaGly 112

Db 214 ---CCCTTCGTGGCGCGCGGTGTGCTGTCGCGGAGGAGGAGGAGGCGCTCGGTGATTG 270

Qy 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132

Db 271 GTGCGGAGGAGCGCTTGTGCTGCTGCTGCGCGCGGTGGAAATTTGTTGGGCG-----321

Qy 133 SerHisThrCysArgSerTyrAspAlaAlaGlyGlyAspHisValLeuValAsp 152

Db 322 -----TGT-----TCTGTGAGGAGGAGGAGGCTATGCGCGG 351

Qy 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrPro 172

Db 352 AAGNNGAGCGCG-----ACGACGCG 372

QY 173 SerSerArg-----AlaHisGlyGluLeuSerAspLeuGluSerAsp----- 186
 Db 373 TCAGACCCCGCCCGCCGAGACCGGACTCGAGCGACGCGGAGTCAACACGAGAGCC 432
 QY 187 -----LeuAlaGlyHisLysThrGlyPro 194
 Db 433 AAGCAGCAATGTGCGCCGGAGTTCGACGACCTCAGCAGCTGCATTTCACGCGGA--- 489
 QY 195 SerLeuProAlaAlaThrProAlaAlaGluLeuLeuValProPro-----AlaHisGlu 212
 Db 490 -----GCCACACGAGGAGCTTCAGGATGATGGCACCCCGCGCGCGCGCAGAG 540
 QY 213 IleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyr 232
 Db 541 ATCGAGGAGTTCTCGCGCTCGGAGAGTCCGAGCGCGCGCTTCGCCCAAGTAC 600
 QY 233 AsnPheAspPheValArgGlyValProLeuAspAlaGly-----GlyArgPheGluTrp 250
 Db 601 AACTTCGACGTGGTGGCGGCGTGGCGCTCGACCGCGCGCGCGCGCGTTCGAATGG 660
 QY 251 AlaProValValSer 255
 Db 661 ACCGCGGTGGCGAGC 675

RESULT 15

US-10-333-006-56
 ; Sequence 56, Application US/10333006
 ; Publication No. US20040019926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
 ; APPLICANT: Peres Bota, Adrian Marius
 ; APPLICANT: Droual, Anne-Marie
 ; APPLICANT: Mironov, Vladimir
 ; APPLICANT: Inz., Dirk
 ; APPLICANT: Hatfield, Yves
 ; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
 ; FILE REFERENCE: 1187-13
 ; CURRENT APPLICATION NUMBER: US/10/333,006
 ; CURRENT FILING DATE: 2003-01-14
 ; PRIOR APPLICATION NUMBER: PCI/IB01/01492
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/218,471
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/241,219
 ; PRIOR FILING DATE: 2000-10-13
 ; NUMBER OF SEQ ID NOS: 59
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 56
 ; LENGTH: 1073
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; US-10-333-006-56

Alignment Scores:
 Pred. No.: 1,62e-14 Length: 1073
 Score: 242.00 Matches: 117
 Percent Similarity: 38.20% Conservative: 27
 Best Local Similarity: 31.03% Mismatches: 90
 Query Match: 18.56% Indels: 144
 DB: 16 Gaps: 19

US-09-993-808B-2 (1-256) x US-10-333-006-56 (1-1073)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaGlu 20
 Db 1 ATGGGGAAG---AAGAGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 57
 QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
 Db 58 GTCGCG-----GGCGTCGTACGCG----- 78
 QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60

Db 79 GCGCCCTCAGCGCGAGGAGGTGTGCGAGCGGAGGAGGTGTGCTTGTGTGGGC 138
 QY 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
 Db 139 CGTGGCGTGGCGGTGGCGAGTGGCGGACGATGGCGAGGGGGGAGTCTATCTGCTG 198
 QY 73 ArgSerArgMetLeu---PheMetAlaProPro-----GlnProGlnPro 86
 Db 199 CGGACGAGAGGTGCTCCCTTGTGGCGCGCGGTGTGTGCTCGCGGAGGAGGAGCG 258
 QY 87 SerValAspSerValProThrProValGluAlaAlaAsp----- 99
 Db 259 CTCGGTGTTCGGTG-----CGGAGGCGGCTTCGCTCGCTCGCGGGCGGTG 309
 QY 100 -----GlyAlaAlaGlyGlnGlnGlyAla----- 107
 Db 310 GAATTTGGTGGCTGTCTGGTGAGGAGGAGGTATGGCCGAGAGGTGATGATGAGCCC 369
 QY 108 -----AlaLeuAlaAlaGlyLeuSerArg---CysSerSerThrAlaSerVal 123
 Db 370 TAGAATTCCTCCGCGCTCGAGTGTCTGATCGCGGCTTCATCTCTTGTGATGATGC 429
 QY 124 AsnLeuGlyLeuGly-----GlyGlnArgGlySerHisThrCys 136
 Db 430 GCGTTGGATGTGTGTTTTCAGGTTCACGCGAGCGAGGATCAGC-ACGAGG 488
 QY 137 Arg-----SerTyrAspAlaAlaGluAlaGlyGlyAspHisVal----- 149
 Db 489 AGAGCTCGTCCGCGACTCCGCTCGCGCGGAGAGGTATCGAGCTCTCTCCACGGG 548
 QY 150 -----LeuValAspValSerAlaAlaSerAsnSerGlySerGlyPro 163
 Db 549 TTCCTGCTTGTCTTGACATGATTATACACCGCGCTCTCTCAATTGAATTATCGCA 608
 QY 164 AspArgGluArgArgGluThrThrProSerSerArg-----AlaHisGlyGluLeu 180
 Db 609 ATTCAATCAGGAGCGCGACGACGCGCTCGAGCCGCGCGCGCGGAGACGCGACTCG 668
 QY 181 SerAspLeuGluSerAsp----- 186
 Db 669 AGCGACGGGAGTCAACACCGAGGAGCCAGAGCAATGTGCGCGGAGTTCGAGACC 728
 QY 187 ---LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaAlaThrProAlaAlaGluLeu 205
 Db 729 TCAGCAGCTGCATTTCACGCGGA-----GCGACGACGAGGAGCTTCAGGATG 776
 QY 206 IleValProPro-----AlaHisGluIleGlnGluPheAlaAlaAlaGluAla 223
 Db 777 ATGGCAGCGCGCGCGCGCGAGATCGAGGAGTTCCTCGCGCTCGCGGAGAGGTCC 836
 QY 224 GlnAlaLysArgPheAlaSerLys----- 231
 Db 837 GAGGCGGAGCGCTTCGCGCGCAAGTGTGTGATCATATTGTGCTCGGTGCTGT 896
 QY 231 ----- 231
 Db 897 GTCGTACATATCGTCGTCTGCTCAAAATCGGCTCGATCGGACATGATGCGCGCANG 956
 QY 232 -----TyrAsnPheAspPheValArgGlyVal 240
 Db 957 GGAGCTGATTACGTGCGCTCTCTCTCAGGTACAACTTCAGCTGTGTCGCGCGGTG 1016
 QY 241 ProLeuAspAlaGly-----GlyArgPheGluTrpAlaProValValSer 255
 Db 1017 CCGCTCGACGCGCGCGCGCGCGGTTCGATGGACCGCGGTGGGCGAGC 1067

Search completed: October 2, 2004, 13:04:07
 Job time : 455 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 12:44:58 ; Search time 3090 Seconds
(without alignments)
3590.879 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256

Sequence: 1 MGXYMKRCGAAGAEVAIVE.....VRGVPLDAGGRFEWAPVWSI 256

Scoring table:

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US0993808/runat_01102004_171726_13707/app_query.fasta_1.455
-DB=GenEmbl -GFFT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptp
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808 @CGN_1_5600 @runat_01102004_171726_13707 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	21	8.2	568	6	AX406682 Sequence
2	17	6.6	1242	6	AX406686 Sequence
3	17	6.5	1266	8	AK103084 Oryza sat
4	17	6.6	135670	8	AF005002 Oryza sat
5	11	4.3	187157	9	AL139384 Human DNA
6	11	4.3	212007	2	AC111297 Rattus no
7	11	4.3	224600	2	AC111313 Rattus no
8	10	3.9	681	6	AX406731 Sequence
9	10	3.9	1073	6	AX406733 Sequence
10	10	3.9	2347	6	AX835276 Sequence
11	10	3.9	2347	6	AX098250 Homo sapi
12	10	3.9	3525	6	AX654623 Sequence
13	10	3.9	10562	1	AE004870 Pseudomon
14	10	3.9	12639	1	AE005072 Halobacte
15	10	3.9	110000	8	AC145127_22 Continuation (23 o
16	10	3.9	151208	2	AF003525 Oryza sat
17	10	3.9	156654	8	AC074283 Oryza sat
18	10	3.9	166660	9	AC008073 Homo sapi
19	10	3.9	187916	2	AP003458 Oryza sat
20	10	3.9	210614	1	AB088224 Streptomy
21	10	3.9	247105	10	AC111653 Rattus no
22	10	3.9	302085	8	AE017054 Oryza sat
23	9	3.5	228	9	AY270716 Homo sapi
24	9	3.5	268	9	AY270729 Homo sapi
25	9	3.5	338	9	Z24649 H. sapiens
26	9	3.5	350	8	HVJ222778 Hordeum v
27	9	3.5	425	9	AY270710 Homo sapi
28	9	3.5	425	9	AY270722 Homo sapi
29	9	3.5	436	6	AR262501 Sequence
30	9	3.5	436	6	AX022310 Sequence
31	9	3.5	436	6	AX030777 Sequence
32	9	3.5	436	6	BD007462 Remedies
33	9	3.5	493	6	AX406678 Sequence
34	9	3.5	688	8	AF205647 Gomphidiu
35	9	3.5	771	6	AR387547 Sequence
36	9	3.5	792	6	AR390079 Sequence
37	9	3.5	975	6	AX573934 Sequence
38	9	3.5	1601	1	AB021882 Streptomy
39	9	3.5	1716	1	AB009289 Streptomy
40	9	3.5	1776	1	AF455754 Bordetell
41	9	3.5	2043	10	AF121907 Mus muscu
42	9	3.5	2113	6	AX834764 Sequence
43	9	3.5	2113	9	AK097418 Homo sapi
44	9	3.5	2163	10	BC017597 Mus muscu
45	9	3.5	2279	8	AK100849 Oryza sat

ALIGNMENTS

RESULT 1

AX406682
 LOCUS AX406682 568 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 5 from Patent WO0228893.
 ACCESSION AX406682
 VERSION AX406682.1 GI:21439630
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Zea.

1
 Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
 and Hatzfeld,Y.

TITLE Plant cyclin-dependent kinase inhibitors
 JOURNAL Patent: WO 0228893-A 5 11-APR-2002;

FEATURES
 CROPDISEGN N V (BE)
 Location/Qualifiers
 1..568
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"

ORIGIN

Alignment Scores:

Pred. No.: 1.02e-07 Length: 568
 Score: 21.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.20% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX406682 (1-568)

QY 225 AlAlysArgPheAlSerIysTyrAnPheAspPheValArgGlyValProLeuAspAla 244
 |||||
 DB 257 GCCAACGCTTGCTTCCAAAGTACAACTTCGACTTCGCGGGCGTGCCTCGACGCC 316
 |||||

QY 245 Gly 245
 ||||
 DB 317 GGC 319

RESULT 2
 AX406686
 LOCUS AX406686 1242 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 9 from Patent WO0228893.
 ACCESSION AX406686
 VERSION AX406686.1 GI:21439634
 KEYWORDS
 SOURCE Zea sativa
 ORGANISM Zea sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

1
 Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
 and Hatzfeld,Y.

TITLE Plant cyclin-dependent kinase inhibitors
 JOURNAL Patent: WO 0228893-A 9 11-APR-2002;

FEATURES
 CROPDISEGN N V (BE)
 Location/Qualifiers
 1..1242
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"

ORIGIN

Alignment Scores:

Pred. No.: 0.000289 Length: 1242
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.64% Indels: 0

DB: 6 Gaps: 0
 US-09-993-808B-2 (1-256) x AX406686 (1-1242)

QY 20 GluValThrGlnValGlyValArgThrArgSerArgSerAlaAlaAla 36
 |||||
 DB 75 GAGGTCCAGCAGGTGGTGGCGTCCGACGAGTCCGCGACGCGCGC 125
 |||||

RESULT 3
 AK103084
 LOCUS AK103084 1266 bp mRNA linear PLN 24-JUL-2003

DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033118G18, full insert sequence.

ACCESSION AK103084

VERSION AK103084.1 GI:32988293

KEYWORDS FLI_CDNA; CAP trapper.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

1
 The Rice Full-Length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-Length cDNA Project Team:
 Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
 Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
 Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
 Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group, Okamoto,Y., Murakami,K.,
 Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
 Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
 Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nishikura,J.,
 Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
 Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
 Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
 Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
 Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
 Saito,R., Sakaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Yoshino,M. and Hayashizaki,Y.

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

2 (bases 1 to 1266)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
 Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
 Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
 Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
 Kanagawa,S., Kato,H., Kawagashira,N., Kawai,J., Kawamata,M.,
 Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
 Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
 Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
 Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
 Namiki,T., Narikawa,R., Nishikura,J., Nishi,K., Nomura,K.,
 Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
 Osato,N., Ota,Y., Ohtsuki,K., Ohtsuki,K., Ohtsuki,K., Ohtsuki,K.,
 Osato,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
 Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
 Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
 Tagawa,A., Takahashi,F., Takaku-Akashira,S., Tanaka,T., Tomaru,A.,
 Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
 Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
 Yoshimura,A.

TITLE

JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT

This clone is one of the 28K full-length cDNA clones from japonica


```

repeat_region      /note="AluX repeat: matches 2. .310 of consensus"
33133. 33489
repeat_region      /note="TIGGER2 repeat: matches 422. .744 of consensus"
33553. 33764
repeat_region      /notes="TIGGER2 repeat: matches 2462. .2690 of consensus"
34943. 35023
misc_feature        /notes="MER81 repeat: matches 7. .88 of consensus"
complement(35414. .35877)
/note="match: GSS: Em:AQ220607"
misc_feature        complement(35464. .35877)
/note="match: GSS: Em:AQ451480"
repeat_region      35690. 35727

```

```

Alignment Scores:
Pred. No.:      726      Length:      187157
Score:          11.00      Matches:      11
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      4.30%      Indels:      0
DB:              9      Gaps:      0

```

US-09-993-808B-2 (1-256) x AL139384 (1-187157)

QY 107 AlaAlaLeuAlaAlaGlyLeuSerArgCysSer 117

Db 68751 GCAGCGCTGCGCCGAGACTGAGCGGTGCTCC 68783

RESULT 6

```

AC111297      212007 bp      DNA      linear      HTG 13-MAY-2003
LOCUS      Rattus norvegicus clone CH230-223J18, *** SEQUENCING IN PROGRESS
DEFINITION      ***. 3 unordered pieces.
AC111297
VERSION      AC111297.5 GI:30578726
KEYWORDS      HTG; PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

REFERENCE

```

AUTHORS
1 (bases 1 to 212007)
Muzny,D,Marle, Metzker,M, Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsebrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Chen,Z., Chu,J.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinu,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Hernandez,B., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,R., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuwa,L., Louisedge,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Manung,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,

```

```

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajda,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G., and Gibbs,R.A.
Direct Submission
2 (bases 1 to 212007)
Worley,K.C.
Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 212007)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:25007665.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLXZ
Center clone name: CH230-223J18
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 192461 bases at least Q40
Consensus quality: 196409 bases at least Q30
Consensus quality: 139024 bases at least Q20
Estimated insert size: 203643; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 208729: contig of 208729 bp in length
* 208730 208829: gap of unknown length
* 208830 210590: contig of 1761 bp in length

```

* 210591 210690: gap of unknown length
 * 210691 212007: contig of 1317 bp in length.

FEATURES

Location/Qualifiers
 1..212007
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-223318"
 misc_feature
 143929..146581
 /note="wgs_contig"
 misc_feature
 154185..156685
 /note="wgs_contig"

ORIGIN

Alignment Scores:
 Pred. No.: 795 Length: 212007
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.30% Indels: 0
 DB: 2 Gaps: 0

US-09-993-808B-2 (1-256) x AC111297 (1-212007)

QY 107 AlaLaLeuAlaAlaGlyLeuSerArgCysSer 117

Db 134431 GCTGCTTTGGCTGCAGGGCTGAGTCGTCCT 134463

RESULT 7

AC111313

LOCUS

DEFINITION

AC111313

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 224600)

Murphy, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyaletchechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Frederick, S., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. B., Hodgson, A., Hogues, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensuewa, L., Loulsegard, H., Iozado, R. J., Iux, K., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, L., Martin, K., Martin, R., Martinez, E.,

Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwaokemele, O., Okwuonu, G., Oiarmpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfanckoch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzoz, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
 Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On May 13, 2003 this sequence version replaced gi:23107935.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GLZY

Center clone name: CH230-243L23

Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 209953 bases at least Q40

Consensus quality: 210980 bases at least Q30

Consensus quality: 211844 bases at least Q20

Estimated insert size: 226554; sum-of-contigs estimation

Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

consists of 3 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

* 1 221276: contig of 221276 bp in length

* 221277 221376: gap of unknown length

* 221377 222686: contig of 1310 bp in length
 * 222687 222786: gap of unknown length
 * 222787 224600: contig of 1814 bp in length.

FEATURES

source
 Location/Qualifiers
 1..224600
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-243L23"

ORIGIN

Alignment Scores:
 Pred. No.: 829 Length: 224600
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.30% Indels: 0
 DB: 2 Gaps: 0

US-09-993-808B-2 (1-256) x AC111313 (1-224600)

QY 107 AlaAlaLeuAlaAlaGlyLeuSerArgCysSer 117

Db 143525 GCTGCTTGGCTGCGGCTGAGTCGCTGCTCT 143557

RESULT 8

AX406731
 LOCUS AX406731 681 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 54 from Patent WO0228893.
 ACCESSION AX406731
 VERSION AX406731.1 GI:21439657

KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.

REFERENCE

1
 AUTHORS Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
 and Hatzfeld, Y.
 TITLE Plant cyclin-dependent kinase inhibitors
 JOURNAL Patent: WO 0228893-A 54 11-APR-2002;
 CROPDESIGN N V (BE)

FEATURES

source
 Location/Qualifiers
 1..681
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="probe or primer"

ORIGIN

Alignment Scores:
 Pred. No.: 74.6 Length: 681
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX406731 (1-681)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246

Db 613 GTGCGCGGCTGCGCTCGACGCGCGGC 642

RESULT 9

AX406733
 LOCUS AX406733 1073 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 56 from Patent WO0228893.
 ACCESSION AX406733
 VERSION AX406733.1 GI:21439658

KEYWORDS
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

1
 AUTHORS Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
 and Hatzfeld, Y.

TITLE Plant cyclin-dependent kinase inhibitors
 JOURNAL Patent: WO 0228893-A 56 11-APR-2002;
 CROPDESIGN N V (BE)

FEATURES

source
 Location/Qualifiers
 1..1073
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"

ORIGIN

Alignment Scores:
 Pred. No.: 104 Length: 1073
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX406733 (1-1073)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246

Db 1005 GTGCGCGGCTGCGCTCGACGCGCGGC 1034

RESULT 10

AX835276/c
 LOCUS AX835276 2347 bp DNA linear PAT 15-DEC-2003
 DEFINITION Sequence 2400 from Patent EPI347046.
 ACCESSION AX835276
 VERSION AX835276.1 GI:39921411

KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
 AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
 Yamamoto, J.I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
 Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
 Masuno, Y.

TITLE Full-length cDNA sequences

JOURNAL Patent: EP 1347046-A 2400 24-SEP-2003;
 Research Association for Biotechnology (JP)

FEATURES

source
 Location/Qualifiers
 1..2347
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 185 Length: 2347
 Score: 100.00 Matches: 10
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX835276 (1-2347)

QY 24 ValValGlyValArgThrArgSerArgSer 33

Db 66 GTGCTAGGCTCGCGACCGCGCTCTCGCTCA 37

RESULT 11

AK098250
 LOCUS AK098250/c 2347 bp mRNA linear PRI 15-JUL-2002
 DEFINITION Homo sapiens cDNA FLJ40931 fis, clone UTERU2007004.

ACCESSION AK098250
 VERSION AK098250.1 GI:21758226
 KEYWORDS oligo capping, fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Mutsaers, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, K., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.

TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2347)
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 COMMENT (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
 source Location/Qualifiers
 1..2347
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UTERU2007004"
 /tissue_type="uterus"
 /clone_lib="UTERU2"
 /note="cloning vector: pME18SFL3"

ORIGIN
 Alignment Scores: Length: 2347
 Pred. No.: 185
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 9 Gaps: 0

US-09-993-808B-2 (1-256) x AK098250 (1-2347)
 QY 24 ValValGlyValArgThrArgSerArgSer 33
 Db 66 GTCGTAGGCGTCCGACCGCTCTCGCTCA 37

RESULT 12
 AX654623
 LOCUS AX654623 3525 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 4493 from Patent WO03000898.
 ACCESSION AX654623
 VERSION AX654623.1 GI:29157437
 KEYWORDS Oryza sativa
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartioideae; Oryzaceae; Oryza.

REFERENCE 1
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M., Katagiri, F., Qian, S., Rao, Y., Whitcham, S., Xie, Z., Zhu, I. and Zou, G.
 TITLE Plant genes involved in defense against pathogens

JOURNAL Patent: WO 03000898-A 4493 03-JAN-2003;
 Syngenta Participations AG (CH)
 FEATURES source Location/Qualifiers
 1..3525
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"

ORIGIN
 Alignment Scores: Length: 3525
 Pred. No.: 249
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX654623 (1-3525)
 QY 28 ArgThrArgSerArgSerAlaAlaAlaThr 37
 Db 206 CGTAGCGCTCGCGCTCTCGCGCGGACG 235

RESULT 13
 AX004870/c
 LOCUS AX004870 10562 bp DNA linear BCT 19-FEB-2003
 DEFINITION Pseudomonas aeruginosa PAO1, section 431 of the complete genome.
 ACCESSION AE004870 AE004091
 VERSION AE004870.1 GI:9950807
 KEYWORDS Pseudomonas aeruginosa PAO1
 SOURCE Pseudomonas aeruginosa PAO1
 ORGANISM Pseudomonas aeruginosa PAO1
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
 REFERENCE 1 (bases 1 to 10562)
 AUTHORS Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Saler, M.H., Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
 TITLE Nature 406 (6799), 959-964 (2000)
 JOURNAL 20437337
 MEDLINE 10964043
 PUBMED
 REFERENCE 2 (bases 1 to 10562)
 AUTHORS Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Saler, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.
 TITLE Direct Submission
 JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA
 REFERENCE 3 (bases 1 to 10562)
 AUTHORS Pseudomonas aeruginosa Community Annotation Project (PseudoCAP)
 REFERENCE Direct Submission
 AUTHORS Submitted (04-FEB-2003) Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada
 JOURNAL

COMMENT
 This represents the February 3, 2003 version of the continually updated, reviewed, Pseudomonas aeruginosa PAO1 genome annotation, from PseudoCAP (see <http://www.pseudomonas.com> for latest updates and links to alternate annotations). PseudoCAP is coordinated by Fiona S.L. Brinkman (Simon Fraser University, Canada) and Robert E.W. Hancock (University of British Columbia, Canada). We welcome

submission through www.pseudomonas.com of any proposed changes.

'Protein name confidence' is used to rate our confidence of the accuracy of the protein name.
 Class 1: Function experimentally demonstrated in *P. aeruginosa*.
 Class 2: Function of highly similar gene experimentally demonstrated in another organism (and gene context consistent in terms of pathways its involved in, if known).
 Class 3: Function proposed based on presence of conserved amino acid motif, structural feature or limited sequence similarity to an experimentally studied gene.
 Class 4: Homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences.

FEATURES

Location/Qualifiers
 1. 10562
 /organism="Pseudomonas aeruginosa PAO1"
 /mol_type="genomic DNA"
 /strain="PAO1"
 /db_xref="taxon:208964"
 complement (79.1617)
 /locus_tag="PA4562"
 /note="synonym: mvn"
 complement (79.1617)
 /locus_tag="PA4562"
 /note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAG07950.1"
 /db_xref="GI:9950808"
 /translation="MMLKSLAVSSITMLSRVLGVFRDITLIRIFGAGIADAFVA
 FKLPLNLRIFAGFASQAPVILABYKQQGEEARTFIAYVSGLLTIVLALVTAG
 ILAAPVIVWTFAPADTPEKFAITDLRTVFPYIILSLSLAGAILNTWNRFSVP
 AFVPTLVNAMIAGFLTPFPYPPVWVLGMAVLGGLLQLPRLKIGMLVLP
 LNRDGSVMKMLPAILGVSQSLIINTIFASFLAAGSVWNYADRLMELPS
 GNLGVALGILLPMLATYSKDRHEYSRLLDWGLLCELVLVPCSLALAEPLTV
 SLFQKFTTVDAMTQRAIVASVGLLILVVKIAPGFIYAQQRITPVIKALFIV
 STQMLNATIGPIQGLALSLIGLACLNAGLWQRLKORLYLPQPGWAKFLVCLV
 AVLMSAVLLATMHWLPFWQGMALRFLRLGLLVVAGLLAYFGLMLLIGFLRLDRSR
 RAVL"
 1901..2176
 /gene="rpsT"
 /locus_tag="PA4563"
 1901..2176
 /gene="rpsT"
 /locus_tag="PA4563"
 /note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
 /codon_start=1
 /transl_table=11
 /product="30S ribosomal protein S20"
 /protein_id="AAG07951.1"
 /db_xref="GI:9950809"
 /translation="MANTPSAKYRAQKERSHNSLSRWVZYIKNVVKAIDAKDL
 EKAQAFTAAVPVIDRMADKGIHKNAKRAKRSLSGHIALSTAA"
 complement (2243..2707)
 /locus_tag="PA4564"
 /note="synonym: creA"
 complement (2243..2707)
 /locus_tag="PA4564"
 /note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAG07952.1"
 /db_xref="GI:9950810"
 /translation="MRKLAGLGLGLLIPALAGAEIEGVSTVFKWGPNDKIWEAF
 DDPKVGVTCTYLSRAKTTGGVKGGLGAEDRAESIACRQVGPIQFSGELKDGEVFKQ

RTSLFKSMQVVRFFDKRNALVYLVLVSDRVIEGSPQNNAVTAIPLMPWPTAR"
 complement (2719..3837)
 /gene="proB"
 /locus_tag="PA4565"
 complement (2719..3837)
 /gene="proB"
 /locus_tag="PA4565"
 /note="Protein name confidence: Class 2 (high similarity to functionally studied protein); Subcellular localization: Cytoplasmic; Subcellular localization confidence: Class 2"
 /codon_start=1
 /transl_table=11
 /product="glutamate 5-kinase"
 /protein_id="AAG07953.1"
 /db_xref="GI:9950811"
 /translation="MRDKVTGARRVWVKI GSALLTADGRGLDRNAVAVVQWVALLHC
 AGIELVSSGAVAGMGRSLGWRSRPSAMHELQAAASVGQMLVQAWESSFALHGLQT
 AOVLTHDDLSDRKRYLNRFADNRPNPDQQLIYEADDPQLDVAAGSAGALGRGSM
 ANIVEADLLVILTRDGMFDADPNPNPDQQLIYEADDPQLDVAAGSAGALGRGSM
 QTKLRAARLAARSGGHTVIVGGRIEVLRLRAGERLCTILLTPDRSRKARKOWLAGH
 QLMRGTLVLDGGAQVSKDQKSLPLPGVAKAVQSGFRGEMVVCVQDGGREVARGLVN
 YSALAEQKILGQPTDAIEALLGYDGPFLVHRDNLVIV"
 complement (3908..5128)
 /gene="obg"
 /locus_tag="PA4566"
 complement (3908..5128)
 /gene="obg"
 /locus_tag="PA4566"
 /note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
 /codon_start=1
 /transl_table=11
 /product="GTP-binding protein Obg"
 /protein_id="AAG07954.1"
 /db_xref="GI:9950812"
 /translation="MKFVDEVSIHVKAGDGGNGLMSFRREKTEKGGNGDGGDGGG
 IYLEADVNTLVYRYTRRFDQARGENGSKDCTAKGDDLILPVPGTVIDANTQ
 EIIGDLEPQORLVAQGGHGLGNTFRKSSINRAPROTTPGKPGEARDLKELKVLA
 DVLGGLPNAGKSTFIRAVSAKGVADYPTFTLVNPGVSVGRYKSFVADIPLGI
 EGAAAGAGIGIRFLHARTRILLHLVMDLPDSDPAEAEVIVRELGRFSFALTER
 ERMVLNKMWDQILDPAEREARKQAVIERLGEWGPYVVISALERDGTALSDQITMRVLD
 ERTLREEDPOYAEELAEELDRITDEARARLQALDDARALRRSLGNAGAVDDDDFDD
 EEDDGDGPRLFYVP"
 complement (5270..5527)
 /gene="rpsM"
 /locus_tag="PA4567"
 complement (5270..5527)
 /gene="rpsM"
 /locus_tag="PA4567"
 /note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
 /codon_start=1
 /transl_table=11
 /product="50S ribosomal protein L27"
 /protein_id="AAG07955.1"
 /db_xref="GI:9950813"
 /translation="MAHKAGSGSTRNGRDSERLGVKLFGGQAVKAGNILLVRQGTGK
 FHAGYGVGLGKDHITLFAKVDGVVVFETKGFAGFKYYSIVAA"
 complement (5551..5862)
 /gene="rplU"
 /locus_tag="PA4568"
 complement (5551..5862)
 /gene="rplU"
 /locus_tag="PA4568"
 /note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
 /codon_start=1
 /transl_table=11
 /product="50S ribosomal protein L21"
 /protein_id="AAG07956.1"
 /db_xref="GI:9950814"
 /translation="MYAVITGQKHQKVTGEELKVEKLDVATGEALDFDRVLLVANG


```
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1617h"
/protein_id="AAG19880.1"
/db_xref="GI:10581092"
/translation="MVDVRYTPADADLWELKRAFERGLGAGTGDSKAKOYADKLIT
ADYARWLQWADCVAESACQVAVESDGLGVFLPESLAFIDAAVLNLEIFVRD
AHRGRGVADELLAAAEVAVAGQSLPLDLRLVLDVDGENERAFAFYERHGFHWGELVAR
DC"
complement(6054..6626)
/gene="VNG1619H"
complement(6054..6626)
/gene="VNG1619H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1619h"
/protein_id="AAG19881.1"
/db_xref="GI:10581093"
/translation="MQSVETAPTTRPATLPAAIGGITAYVVGAYVTYLMQSDAVGAAL
QSFNAVIALGGDPTITWAGWLFYNAAHAFSTFAVGSRLSRLLIDGNAPLVLL
LIPAGLVVAGTAGHDGTTPRADRLRGVRVTAGYAAVNLGVALFRFTAGTATVH
VEYALGVLVAGIGYPAVGAAGALAAVTR"
complement(6675..7130)
/gene="VNG1621H"
complement(6675..7130)
/gene="VNG1621H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1621h"
/protein_id="AAG19882.1"
/db_xref="GI:10581094"
/translation="MTDTVAHLEPPDRDAFLAPGCTGVLSFGREDAGPPHSLPVSYG
YDPTSFRRSLYSDDTKPDPETPLDVSFVHDEADHGWCSVVAHGRLOPTDADA
ISTTALSALRRVDIPLLDTFDRPLVETEFLEFLHDPDAITARADAPPTT"
7129..8634
/gene="rfcB"
/note="synonym: VNG1622G"
7129..8634
/gene="rfcB"
/note="rfcB"
/codon_start=1
/transl_table=11
/product="replication factor C large subunit"
/protein_id="AAG19883.1"
/db_xref="GI:10581095"
/translation="MPCVATERGKTVAGAPQHLTRTPAYGVSMVDWTEKYRPASIS
EYRGNTARDALAEWAETWPDHREAVVHGSFGIKGTSAAHALANDAGWDVVELNASD
QRTADVVERVAGEAASGTLTGSGGGRKVLLEADNLGNIDRGSSAALTRLVDDAP
QPVLIVANYYEMSSISRACTIEFRDVSRSKRSIVPLRDVCRREDVTYEDLAAIA
EQAGDLRSANDLQALASQDRYTLTADDVVMGERDTEGVFDYLDVVIATHSAREALQ
AAVDDETDDLSSWADNPKDYRGGELADAYEFSLNADWILGRVATQNYAWRYA
TNNVAGAAARLHHGGWTRYGPSPYWRKLGSSRATREKRDYVARSHIAETAGCSMAT
ARNDLPELRLVTHCKNDELTVAMAAAYELDTHEVAFVTGSGETTNNKVASIVADAE
RRDVAVDHSEGFAGAVREDNTDSDAEDTDDGDTGDSQSGLDSEFF"
complement(8631..9464)
/gene="ccp"
/note="synonym: VNG1623G"
complement(8631..9464)
/gene="ccp"
/note="ccp"
/codon_start=1
/transl_table=11
/product="cytochrome aa3 controlling protein"
/protein_id="AAG19884.1"
/db_xref="GI:10581096"
/translation="MARRSQALAAVTAFLTFLALLIGETFANVSGSGATCGLQWPCSG
QFLPGLPVNDFLEQHRSVAMVVGFI LGTGAAYRNFDARDIKIAGVLAIVLLPIQ
VILGGTVTFPGVQPLTQAVHLLAAMAILGTVTYALRMRDAAIGILGRVTA
VLAALAVLPEFAVSRGTFTYVGSWITQIIHHTFELALEFTAAFTAFVWLSRRNTRSA
AATGGVALLSTAQVLLGVLPFGPTIQLAYYALAVVIAALLVATARIHAHAPPTAAT
```

```
gene A"
9595..11847
/feature="mdh"
/note="synonym: VNG1624G"

Alignment Scores:
Pred. No.: 635 Length: 12639
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 1 Gaps: 0

US-09-993-808B-2 (1-256) x AE005072 (1-12639)

QY 32 ArgSerAlaAlaThrGlyValAla 41
|||
Db 8819 CGCAGCGCGCGCCGCCACCGCGGTGCGG 8790
|||

RESULT 15
AC145127_22/c
WPCOMMENT
Sequence split into 24 fragments LOCUS AC145127 Accession AC145127
Fragment Name Begin End
AC145127_00 1 110000
AC145127_01 100001 210000
AC145127_02 200001 310000
AC145127_03 300001 410000
AC145127_04 400001 510000
AC145127_05 500001 610000
AC145127_06 600001 710000
AC145127_07 700001 810000
AC145127_08 800001 910000
AC145127_09 900001 1010000
AC145127_10 1000001 1110000
AC145127_11 1100001 1210000
AC145127_12 1200001 1310000
AC145127_13 1300001 1410000
AC145127_14 1400001 1510000
AC145127_15 1500001 1610000
AC145127_16 1600001 1710000
AC145127_17 1700001 1810000
AC145127_18 1800001 1910000
AC145127_19 1900001 2010000
AC145127_20 2000001 2110000
AC145127_21 2100001 2210000
AC145127_22 2200001 2310000
AC145127_23 2300001 2331000

Continuation (23 of 24) of AC145127 from base 2200001 (AC145127 Oryza sativa (japonica cv

Alignment Scores:
Pred. No.: 3.1e+03 Length: 110000
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 8 Gaps: 0

US-09-993-808B-2 (1-256) x AC145127_22 (1-110000)

QY 28 ArgThrArgSerArgSerAlaAlaThr 37
|||
Db 31035 CGTACGCGTCCGCTCTGCGCGGCGGCGACG 31006
|||

Search completed: October 2, 2004, 14:04:43
Job time : 3251 secs
```

This Page Blank (uspto)

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 12:50:51 ; Search time 2465 Seconds
(without alignments)
3101.306 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVLPDAGGRFEPWVSI 256

Scoring table:
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=franc+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPRO.spool_p/US0993808/runat_01102004_171727_13719/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808@cgn_1_1_5180@runat_01102004_171727_13719 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FPGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	127	49.6	526	14	CD444277	CD444277 EL01N0438	
C	2	65	25.4	753	29	CG729673	CG729673 OGBY357H
C	3	65	25.4	751	29	CG324885	CG324885 OXKDN04TV
C	4	65	25.4	825	29	CG265472	CG265472 OQ2CM60TH
C	5	65	25.4	890	29	CG218373	CG218373 OGLD110TV
C	6	65	25.4	917	29	CG265481	CG265481 OQ2CM60TV
C	7	65	25.4	986	28	CC002416	CC002416 PUGKB53TB
C	8	65	25.4	987	28	CC002419	CC002419 PUGKB53TD
C	9	65	25.4	1010	28	CC390569	CC390569 PUHBT76TB
C	10	61	23.8	707	11	AY110319	AY110319 Zea mays
C	11	59	23.0	547	28	BH777311	BH777311 f2mb013f0
C	12	47	18.4	645	29	CG136784	CG136784 PU1DC25TB
C	13	43	16.8	488	28	BH784474	BH784474 f2mb013f0
C	14	43	16.8	666	29	CC610974	CC610974 OGVFR67TV
C	15	43	16.8	838	28	CC440801	CC440801 PUHLN53TB
C	16	43	16.8	869	28	BZ743686	BZ743686 OGFAC93TC
C	17	42	16.4	225	14	CD957606	CD957606 SKL_144 G
C	18	41	16.0	644	29	CC729680	CC729680 OGBY35TV
C	19	37	14.5	655	13	CA104638	CA104638 SCJFHR103
C	20	36	14.1	659	13	CA074449	CA074449 SCEZAM108
C	21	34	13.3	510	14	CB929335	CB929335 ABA1_41 H
C	22	29	11.3	413	14	CD951880	CD951880 SAZ_175 G
C	23	29	11.3	657	28	BZ743690	BZ743690 OGFAC93TM
C	24	28	10.9	446	14	CF632445	CF632445 zmrvs48 0
C	25	28	10.9	531	14	CA297189	CA297189 SCUTSD208
C	26	27	10.5	563	13	CA108367	CA108367 SCSFHR104
C	27	27	10.5	584	13	CA108352	CA108352 SCSFHR104
C	28	27	10.5	705	13	CA106521	CA106521 SCOGHR101
C	29	27	10.5	885	29	CG324875	CG324875 OXKDN04TH
C	30	25	9.8	405	28	BH879152	BH879152 h886e11.b
C	31	25	9.8	585	13	CA104568	CA104568 SCJFHR103
C	32	25	9.8	661	14	CA256119	CA256119 SCJLFL418
C	33	25	9.8	741	29	CC683397	CC683397 OGUHL76TV
C	34	25	9.8	790	28	CC390573	CC390573 PUHBT76TD
C	35	25	9.8	835	29	CG218361	CG218361 OGLD110TH
C	36	25	9.8	887	29	CG223296	CG223296 OGMAL17TV
C	37	21	8.2	550	12	BG267004	BG267004 1000110B0
C	38	21	8.2	568	9	AI737717	AI737717 605040C07
C	39	21	8.2	642	13	CA104689	CA104689 SCJFHR103
C	40	21	8.2	1197	11	AY108168	AY108168 Zea mays
C	41	17	6.6	697	14	CB646543	CB646543 OSJNEB09B
C	42	17	6.6	708	28	BZ404098	BZ404098 OGACT23TM
C	43	17	6.6	753	14	CB646893	CB646893 OSJNEB09J
C	44	17	6.6	787	28	BZ404090	BZ404090 OGACT23TC
C	45	17	6.6	795	29	CG339703	CG339703 OG1DH12TH

ALIGNMENTS

RESULT 1
CD444277
LOCUS EL01N0438A02.b Endospem_4 Zea mays mRNA linear EST 03-JUN-2003
DEFINITION
ACCESSION CD444277
VERSION CD444277.1 GI:31359920
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 526)

AUTHORS Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and Messing, J.
 TITLE Sequencing of the maize endosperm ESTs
 JOURNAL Unpublished (2002)
 COMMENT Contact: Lai, Jinsheng
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T3.

FEATURES
 source
 1..526
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="W22"
 /db_xref="taxon:4577"
 /tissue_type="Endosperm of 7-23DAP"
 /clone_lib="Endosperm 4"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
 Alignment Scores:
 Pred. No.: 3.99e-96 Length: 526
 Score: 127.00 Matches: 127
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 49.61% Indels: 0
 DB: 14 Gaps: 0

US-09-993-808B-2 (1-256) x CD444277 (1-526)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGlnValAlaValGlu 20
 Db 146 ATGGGGAAGTACATGCGCAAGTGCAGGGCGCGCAGGCGCGAGTCCGCCGCGTCGAG 205

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
 Db 206 GTTAGCAGGTGTCGGCGTCCGAGAGGTCCAGGTCCGGCGCGCGCGCGCGGTGTC 265

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
 Db 266 GCGAAGGTGCGCCCGAGGAGGAGGCGCGCGGGGGAGCTCTCCCGCGGTGAGC 325

QY 61 AlaGlyGlyAspGlySerCysTyrIleHisLeuArgSerArgValLeuPheMetAla 80
 Db 326 GCTGTGGGGAGCGGGAAGTGTACATCCACCTCGTAGCGCATGCTGTTCATGGCA 385

QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
 Db 386 CCGCCTCAGCCGAGCGCGTGGTTGACTCGGTTCGACCCCGTGGAGGCTGCTGATGGC 445

QY 101 AlaAlaGlyGlnGlnAlaAlaLeuAlaAlaGlyLeuSerArgCysSerThrAla 120
 Db 446 GCTGCAGGACAGCAGGGCGCGGCTGCTGCGGGCTCTCGCGTTGCTCCAGCAGCGCG 505

QY 121 SerSerValAsnLeuGlyLeu 127
 Db 506 TCGTCTGTGAACCTTGGGCTTG 526

RESULT 2
 CC729673/C
 LOCUS CC729673 753 bp DNA linear GSS 23-JUN-2003
 DEFINITION OGUB35TH_ZM_0.7_1.5_KB Zea mays genomic clone ZMMEMa0404F21,
 genomic survey sequence.
 ACCESSION CC729673
 VERSION CC729673.1 GI:32148606
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGUB35TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: PR
 Class: sheared ends.
 Location/Qualifiers
 1..753
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMEMa0404F21"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Alignment Scores:
 Pred. No.: 4.59e-44 Length: 753
 Score: 65.00 Matches: 65
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 25.39% Indels: 0
 DB: 29 Gaps: 0

US-09-993-808B-2 (1-256) x CC729673 (1-753)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluSerAspLeuGluSerAsp 186
 Db 410 AGGAGAGAGACACGCCATCGAGCGCGCGAGCTCAGCGATCTGGAGTCGGAT 351

QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
 Db 350 CTGGCGGGGCAAGACTGGCCGCTCGTACCGCGGCAACGCCGCTCGGAGCTGATC 291

QY 207 ValProProAlaHisGluIleGlnGluPhePheAlaAlaAlaGluAlaGlnAlaLys 226
 Db 290 GTGCCGCCACACACAGATCCAGAGTTCTTCGCCGCCGCCGAGCGCCGCCCAAG 231

QY 227 ArgPheAlaSerLys 231
 Db 230 CGCTTGTCTTCAAG 216

RESULT 3
 CC729673/C
 LOCUS CC729673 791 bp DNA linear GSS 26-AUG-2003
 DEFINITION OGXD04TV_ZM_0.7_1.5_KB Zea mays genomic clone ZMMEMa0661B08,
 genomic survey sequence.
 ACCESSION CC729673
 VERSION CC729673.1 GI:34242151
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 791)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics

Alignment Scores:			
Pred. No.:	5.64e-44	Length:	917
Score:	65.00	Matches:	65
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	25.39%	Indels:	0
DB:	29	Gaps:	0
US-09-993-808B-2 (1-256) x CG265481 (1-917)			
Qy	167	AtgAtgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp	186
Db	262	AGAGAGAGACGACGCCATTCGAGCCGGCGGCGACCGGAGCTCAGCGATCTGGAGTCGGAT	321
Qy	187	LeuAlaGlyHisLysThrGlyProSerLeuProLalaThrProAlaAlaGluLeuLeu	206
Db	322	CTGGCGGGGACACAGACTGGCCGCTGCTACCGGCGGCACACGCGGCTGCGGAGCTGATC	381
Qy	207	ValProProAlaHisGluLeuGlnGluPheAlaAlaGluAlaAlaGluAlaAlaLys	226
Db	382	GTGCCGCCAGACACAGATCCAGAGTCTTCGCCGCCGCGCGGCGGCCAGGCGCAAG	441
Qy	227	AtgPheAlaSerLys 231	
Db	442	CGCTTGTCTTCCAG 456	
<div> <div> <div>RESULT 7</div> <div>CC002416</div> <div>LOCUS</div> <div>DEFINITION</div> <div>ACCESSION</div> <div>VERSION</div> <div>KEYWORDS</div> <div>SOURCE</div> <div>ORGANISM</div> </div> <div> <div>PUGXB53TB</div> <div>ZM_0.6_1.0_KB</div> <div>Zea mays genomic clone ZMWBta402110,</div> <div>genomic survey sequence.</div> <div>CC002416</div> <div>CC002416.1</div> <div>GI:29380976</div> <div>GSS.</div> <div>Zea mays</div> <div>Zea mays</div> <div>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.</div> <div>1 (bases 1 to 986)</div> <div>Whitelaw,C.A.; Quackenbush,J.; Van Aken,S.; Utterback,T.; Resnick,A.; Fraser,C.M.; Yuan,Y.; San Miguel,P., Ma,J. and Bennettzen,J.</div> <div>Maize Genomics Consortium</div> <div>Unpublished (2003)</div> <div>Other GSSs: PUGXB53TD</div> <div>Contact: Cathy Whitelaw</div> <div>TIGR</div> <div>9712 Medical Center Drive, Rockville, MD 20850, USA</div> <div>Tel: 301-838-5843</div> <div>Fax: 301-838-0208</div> <div>Email: whitelaw@tigr.org</div> <div>Seq primer: TR</div> <div>Class: sheared ends.</div> </div> </div> <div> <div> <div>FEATURES</div> <div>source</div> </div> <div> <div>Location/Qualifiers</div> <div>1..986</div> <div>/organism="Zea mays"</div> <div>/mol_type="genomic DNA"</div> <div>/strain="B73"</div> <div>/db_xref="taxon:4577"</div> <div>/clone="ZMWBta402110"</div> <div>/clone_lib="ZM_0.6_1.0_KB"</div> <div>/notes="vector: PCR4-TOPO; Site.1: EcoRI; 0.6-1.0 kb high Cot selected genomic DNA library"</div> </div> </div> <div> <div> <div>ORIGIN</div> </div> <div> <div>Alignment Scores:</div> <div>Pred. No.:</div> <div>Score:</div> <div>Percent Similarity:</div> <div>Best Local Similarity:</div> <div>Query Match:</div> <div>DB:</div> </div> <div> <div>6.09e-44</div> <div>65.00</div> <div>100.00%</div> <div>100.00%</div> <div>25.39%</div> <div>29</div> </div> <div> <div>Length:</div> <div>Matches:</div> <div>Conservative:</div> <div>Mismatches:</div> <div>Indels:</div> <div>Gaps:</div> </div> <div> <div>986</div> <div>65</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> </div> </div>			

```

DB:                28                Gaps:                0
US-09-993-808B-2 (1-256) x CC002416 (1-986)

Qy      167 ArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
Db      346 AGGCGAGAGACGACGCCATCGAGCGCGCGCACCGCGAGCTCGAGTCGGAT 405

Qy      187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
Db      406 CTGCGGGGGCACAAGACTGGCCCTCGCTACCGCGCGCAACGCCGGCTCGCGAGCTGATC 465

Qy      207 ValProProAlaHisGluLeuGlnGluPheAlaAlaGluAlaGlnAlaLys 226
Db      466 GTGCGCGCAGCAGCAGATCCAGAGTTCCTTTCGCCGCCGCGAGCGGCCAGGCCAAG 525

Qy      227 ArgPheAlaSerLys 231
Db      526 CGCTTTGCTTCCAAG 540

RESULT 8
LOCUS      CC002419/c
DEFINITION PUGKB53TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBtra402110,
ACCESSION  CC002419
VERSION     CC002419.1 GI:29380979
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1 (bases 1 to 987)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
TITLE       Maize Genomics Consortium
JOURNAL     Unpublished (2003)
COMMENT     Other_GSSs: PUGKB53TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
FEATURES    source
             Location/Qualifiers
               1..987
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBtra402110"
                /clone_lib="ZM 0.6 1.0 KB"
                /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                COT selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.:      6,1e-44                Length:      987
Score:          65.00                Matches:      65
Percent Similarity: 100.00%          Conservative: 0
Best Local Similarity: 100.00%        Mismatches:  0
Query Match:    25.39%                Indels:      0
DB:             28                Gaps:            0

US-09-993-808B-2 (1-256) x CC002419 (1-987)

Qy      167 ArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
Db      640 AGGCGAGAGACGACGCCATCGAGCGCGCGCACCGCGAGCTCGAGTCGGAT 581

```

```

Qy      187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
Db      580 CTGCGGGGGCACAAGACTGGCCCTCGCTACCGCGCGCAACGCCGGCTCGCGAGCTGATC 521

Qy      207 ValProProAlaHisGluLeuGlnGluPheAlaAlaGluAlaGlnAlaLys 226
Db      520 GTGCGCGCAGCAGCAGATCCAGAGTTCCTTTCGCCGCCGCGAGCGGCCAGGCCAAG 461

Qy      227 ArgPheAlaSerLys 231
Db      460 CGCTTTGCTTCCAAG 446

RESULT 9
LOCUS      CC390569
DEFINITION PUHBT76TB ZM 0.6 1.0 KB Zea mays genomic clone ZMMBtra415M08,
ACCESSION  CC390569
VERSION     CC390569.1 GI:30870659
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1 (bases 1 to 1010)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
TITLE       Maize Genomics Consortium
JOURNAL     Unpublished (2003)
COMMENT     Other_GSSs: PUHBT76TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES    source
             Location/Qualifiers
               1..1010
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBtra415M08"
                /clone_lib="ZM 0.6 1.0 KB"
                /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                COT selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.:      6,24e-44                Length:      1010
Score:          65.00                Matches:      65
Percent Similarity: 100.00%          Conservative: 0
Best Local Similarity: 100.00%        Mismatches:  0
Query Match:    25.39%                Indels:      0
DB:             28                Gaps:            0

US-09-993-808B-2 (1-256) x CC390569 (1-1010)

Qy      167 ArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
Db      177 AGGCGAGAGACGACGCCATCGAGCGCGCGCACCGCGAGCTCGAGTCGGAT 236

Qy      187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
Db      237 CTGCGGGGGCACAAGACTGGCCCTCGCTACCGCGCGCAACGCCGGCTCGCGAGCTGATC 296

Qy      207 ValProProAlaHisGluLeuGlnGluPheAlaAlaGluAlaGlnAlaLys 226

```


genomic survey sequence.

ACCESSION
CG136784
VERSION
CG136784.1 GI:34026526
KEYWORDS
GSS.
SOURCE
Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS

1 (bases 1 to 645)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennettzen,J.

TITLE
Maize Genomics Consortium

JOURNAL

COMMENT

Unpublished (2003)

Other_GSSs: PUIDC25TD

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

Location/Qualifiers

1..645

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBTa0556E01"

/clone_lib="ZM 0.6-1.0 KB"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high

ORIGIN

Alignment Scores:

Pred. No.: 4.56e-29 Length: 645

Score: 47.00 Matches: 47

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 18.36% Indels: 0

DB: 29 Gaps: 0

US-09-993-808B-2 (1-256) x CG136784 (1-645)

Qy 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186

Db 503 AGCGAGAGACGACGCCATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 562

Qy 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaThrProAlaAlaGluLeuIle 206

Db 563 CTGCGGGGACAAAGATGGCCCTCGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 622

Qy 207 ValProProAlaHisGluIle 213

Db 623 GTGCCGCGCAGCACACGAGATC 643

RESULT 13

BH784474

LOCUS

DEFINITION

fzmb013f023f06k0 fzmb filtered library Zea mays genomic clone

BH784474 488 bp DNA linear GSS 28-MAR-2002

fzmb013f023f06 5', genomic survey sequence.

ACCESSION

VERSION

BH784474.1 GI:19788171

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 488)

Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.

GeneThresher methylation filtered genomic sequences from maize

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Contact: Bedell JA

Orion Genomics, LLC

4041 Forest Park Ave, St. Louis, MO 63108, USA

Tel: 314 615 6979

Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: fzmb013f023 row: f column: 06

Seq primer: SK reverse

Class: shotgun

High quality sequence stop: 488.

Location/Qualifiers

1..488

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="Mol7"

/db_xref="taxon:4577"

/clone="fzmb013f023f06"

/clone_lib="fzmb filtered library"

/note="Organ: leaf; Vector: pBCKS(-); Site 1: HincII; DNA

prepared from purified nuclei was randomly sheared,

end-repaired, size fractionated to enrich for the 0.5 to

5 kb fraction, ligated into HincII-digested pBCKS(-)

vector and electroporated into E. coli cells."

ORIGIN

Alignment Scores:

Pred. NO.: 7.59e-26 Length: 488

Score: 43.00 Matches: 43

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 16.80% Indels: 0

DB: 28 Gaps: 0

US-09-993-808B-2 (1-256) x BH784474 (1-488)

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyValAlaGlyValAlaValGlu 20

Db 323 ATGGGAAGTACATGCCAAGTCCAGGGGGCGCGCGCGCGCGCGCGCGCGCGAT 382

Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyVal 40

Db 383 GTTACGCGAGTCTCGCGGTCGCGCACGAGTCCAGTCCGCGCGCGCGCGGTC 442

Qy 41 AlaLysVal 43

Db 443 GCGAAGGTC 451

RESULT 14

CC610974/c

LOCUS

DEFINITION

OGVFR67TV ZM 0.7-1.5_KB Zea mays genomic clone ZMMBa0551L14,

genomic survey sequence.

ACCESSION

CC610974

VERSION

CC610974.1 GI:31972395

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 666)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other_GSSs: OGVFR67TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES

source
1..666
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0551L14"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBGSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 1.05e-25 Length: 666
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.80% Indels: 0
DB: 29 Gaps: 0
US-09-993-808B-2 (1-256) x CC610974 (1-666)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyValAlaGluValAlaValGlu 20
Db 467 ATGGGAAAGTACATCGCAAGTGCAGGGGGCGCGAGGTCGCCGCCGCGCGAG 408
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 407 GTTACGACGTCGTGGCGTCCGCACAGGTCACAGTCCGCGCGCGACCGCGCGTGC 348
QY 41 AlaLysVal 43
Db 347 GCGAAGGTC 339

RESULT 15

CC440801/c
LOCUS CC440801 838 bp DNA linear GSS 20-MAY-2003
DEFINITION PUHLN53TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTA478J10,
genomic survey sequence.

ACCESSION

VERSION CC440801.1 GI:30942238
KEYWORDS GSS.

SOURCE

Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 838)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.

REFERENCE

AUTHORS Maize Genomics Consortium
TITLE Unpublished (2003)
JOURNAL Other GSSs: PUHLN53TD
COMMENT Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES

source
1..838
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA478J10"

/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pC34-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 1.34e-25 Length: 838
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.80% Indels: 0
DB: 28 Gaps: 0
US-09-993-808B-2 (1-256) x CC440801 (1-838)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyValAlaGluValAlaValGlu 20
Db 390 ATGGGAAAGTACATCGCAAGTGCAGGGGGCGCGAGGTCGCCGCCGCGCGAG 331
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 330 GTTACGACGTCGTGGCGTCCGCACAGGTCACAGTCCGCGCGCGACCGCGCGTGC 271
QY 41 AlaLysVal 43
Db 270 GCGAAGGTC 262

Search completed: October 2, 2004, 14:43:24
Job time : 2470 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 11:51:15 ; Search time 366 Seconds

(without alignments)

2971.417 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256

Sequence: 1 MGKYMRCRGAAGAEEVAABE.....VRGVPLDAGGFEWAPVWSI 256

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745146

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q/cen2_1/UsPto_spool_p/US09993808/runat_01102004_171726_13699/app_query.fasta_1.455
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=us09993808 @CEN 1 1 708 @runat_01102004_171726_13699 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04.*

1: Geneseqn13908.*
2: Geneseqn13908.*
3: Geneseqn20000.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	256	100.0	1372	7	ABV74603 Maize CKI
2	101	39.5	723	3	AAA95281
3	21	8.2	568	6	ABK93954 Corn zmic
4	21	8.2	572	3	AA02390 Corn cycl
5	21	8.2	572	3	AAA95276
6	21	8.2	639	3	AAA95280 Corn cycl
7	17	6.6	841	7	ABV74605
8	17	6.6	1242	6	ABK93958 DNA encod

9	10	3.9	681	6	ABK93981	Abk93981 Full leng
10	10	3.9	1073	6	ABK93982	Abk93982 Rice OsIC
11	10	3.9	1539	7	ACA42627	Aca42627 Prokaryot
12	10	3.9	3525	7	ADA71170	Ada71170 Rice gene
13	9	3.5	204	3	AAA95286	Aaa95286 Rice cycl
14	9	3.5	204	3	AA02400	Aa02400 Rice Cycl
15	9	3.5	334	6	ABL81995	AbL81995 Human ova
16	9	3.5	392	4	AKS6036	Aks6036 Human ihm
17	9	3.5	414	8	ACH30301	Ach30301 Human tes
18	9	3.5	436	2	AAV38676	Aav38676 Mus muscu
19	9	3.5	484	7	ACA35095	Aca35095 Prokaryot
20	9	3.5	493	6	ABK93950	Abk93950 Rice OsIC
21	9	3.5	499	8	ACL13592	ACL13592 DNA clone
22	9	3.5	500	8	ACL25923	ACL25923 DNA clone
23	9	3.5	501	8	ACL25927	ACL25927 DNA clone
24	9	3.5	529	8	ACL25925	ACL25925 DNA clone
25	9	3.5	530	8	ACL25926	ACL25926 DNA clone
26	9	3.5	571	8	ACL13581	ACL13581 DNA clone
27	9	3.5	571	8	ACL13580	ACL13580 DNA clone
28	9	3.5	578	8	ACL13594	ACL13594 DNA clone
29	9	3.5	597	8	ACL25923	ACL25923 DNA clone
30	9	3.5	630	8	ACL13582	ACL13582 DNA clone
31	9	3.5	631	8	ACL13590	ACL13590 DNA clone
32	9	3.5	632	8	ACL13589	ACL13589 DNA clone
33	9	3.5	634	8	ACL13591	ACL13591 DNA clone
34	9	3.5	634	8	ACL13595	ACL13595 DNA clone
35	9	3.5	634	8	ACL13584	ACL13584 DNA clone
36	9	3.5	635	8	ACL25924	ACL25924 DNA clone
37	9	3.5	646	8	ACL13593	ACL13593 DNA clone
38	9	3.5	663	8	ACL25922	ACL25922 DNA clone
39	9	3.5	663	8	ACL13579	ACL13579 DNA clone
40	9	3.5	667	8	ACL13596	ACL13596 DNA clone
41	9	3.5	675	8	ACL25928	ACL25928 DNA clone
42	9	3.5	677	8	ACL13583	ACL13583 DNA clone
43	9	3.5	677	8	ACL13577	ACL13577 DNA clone
44	9	3.5	822	3	AAA45372	Aaa45372 Human sec
45	9	3.5	975	7	ABZ66676	Abz66676 Orthosomy

ALIGNMENTS

RESULT 1

ABV74603

ID ABV74603 standard; DNA; 1372 BP.

XX AC ABV74603;

XX DT 21-FEB-2003 (first entry)

XX DE Maize CKI_B coding sequence.

XX KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;

XX KW crop yield; root size; plant growth; tassel size; ear size;

XX KW male sterility; endoreduplication; gene; ds.

XX OS Zea mays.

XX FH Key

CDS Location/Qualifiers

FT 134..904

FT /*tag= a

FT /product= "CKI_B"

XX PN WO200281623-A2.

XX PD 17-OCT-2002.

XX PF 06-NOV-2001; 2001WO-US044038.

XX PR 07-NOV-2000; 2000US-0246349P.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PA (ARIZ-) ARIZONA BOARD OF REGENTS.

PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
 XX WPI; 2003-058511/05.
 DR P-PSDB; ABB98757.
 XX

Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
 PT involved in cell cycle regulation, and useful for altering cell cycle
 PT protein content, cell cycle progression, cell number and composition of
 PT plants.
 XX

Claim 1; Page 64-65; 69pp; English.

XX The present sequence is the coding sequence (I) for maize cyclin-
 CC dependent kinase inhibitor (CKI). CKI B. (I) is useful for modulating the
 CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
 CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
 CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
 CC preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain
 XX

SQ Sequence 1372 BP; 292 A; 396 C; 442 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 4.01e-226 Length: 1372
 Score: 256.00 Matches: 256
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-09-993-808B-2 (1-256) x ABV74603 (1-1372)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
 DB 134 ATGGGGAAGTACATGCGCAAGTGCAGGGGCGCGCGAGGAGGTCGCGCGCGTCGAG 193

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
 DB 194 GTTACGCGAGTCTGTCGGGTCGCGAGGTCAGGTCGCGCGCGGCGGCGCGGTCGTC 253

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
 DB 254 GCGAAGGTGCGCCCGAGGAGGAGGAGGCGCGCGCGGAGGCGCGCGCGCGCGCGGAGC 313

QY 61 AlaGlyGlyAspGlyCysCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
 DB 314 GCTGTGGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 373

QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
 DB 374 CCGCTCAGCGCGAGCGGTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 433

QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
 DB 434 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 493

QY 121 SerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTrpAsp 140
 DB 494 TCGTGGTGAATTTGGGGGTTGGGGGTCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 553

QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
 DB 554 GCTGACAGAGGTCGCGGGGATCAGTCTCTGTGGATGTCCTCGGCGGCGAGCACTCGGG 613

QY 161 SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
 DB 614 AGCGGCCAGACCGCGAGCGAGACCGACCGACCGACCGACCGACCGACCGACCGACGTC 673

QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
 DB 674 AGCGATCTGGAGTCGATCTGGCGGGGCAACAGACTGGCGGTCGCTACCGCGGCAACG 733

QY 201 ProAlaAlaGluLeuValProProAlaHisGluLeuGlnGluPhePheAlaAlaAla 220
 DB 734 CCGGCTCGGAGTGTGTCGCGCGGACACAGAGATCCAGGAGTTCCTCGCGCGCGCGCC 793

QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
 DB 794 GAGCGCGCCAGGCGGCGGCTTTGCTTCCAAAGTACAACTTCGACTTCGTCGCGCGGCGTG 853

QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
 DB 854 CCCCTCGACGCGCGCGCGGTCGAGTGGCGCGCGGTCGTCAGCATC 901

RESULT 2
 ID AAA95281 standard; cDNA; 723 BP.
 XX AAA95281;
 AC AAA95281;
 XX 17-JAN-2001 (first entry)
 DT
 XX Corn cyclin-dependent kinase inhibitor coding sequence #3.
 DE
 XX Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 KW cell growth; herbicide; ss.
 XX
 OS Zea mays.
 XX
 XX Key Location/Qualifiers
 XX CDS 3..380
 XX /*tag= a
 XX /product= "CDKI"
 XX FT
 XX FT
 XX FT
 XX PN WO200060087-A2.
 XX PD 12-OCT-2000.
 XX
 XX PF 06-APR-2000; 2000WO-US009106.
 XX PR 07-APR-1999; 99US-0128192P.
 XX
 XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX PI Klein TM, Weng Z, Cahoon RE;
 XX
 XX WPI; 2000-679375/66.
 DR P-PSDB; AAB26250.
 DR
 XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.
 XX
 XX Claim 2; Page 43-44; 58pp; English.
 XX
 CC The present sequence is the coding sequence for the corn cyclin-dependent
 CC kinase inhibitor (CDKI). It was isolated by searching a contig comprising
 CC cDNA from corn roots and ear leaf sheath for sequences similar to those
 CC encoding the CDKI from *Chenopodium rubrum*, *Caenorhabditis elegans* and
 CC *Arabidopsis thaliana*. CDKI is involved in the cell cycle, and may promote
 CC or inhibit cell division and growth. The coding sequence and the protein
 CC it encodes are useful in the production of transgenic plants which
 CC produce increased or decreased amounts of the CDKI protein, in the

CC identification of herbicides, in genetic and physical mapping and in the
 XX isolation of the CDKI gene in other organisms
 SQ Sequence 723 BP; 160 A; 202 C; 210 G; 151 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.59e-83 Length: 723
 Score: 101.00 Matches: 101
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 39.45% Indels: 0
 DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAA95281 (1-723)

QY 156 AlaSerAsnSerGlySerGlyProAspArgGluThrThrProSerSerArg 175
 Db 75 GCAGCAACTCCCGGAGCGCCAGACCGGAGAGAGAGAGAGAGAGAGAGCGG 134

QY 176 AlaHisGlyGluSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSer 195
 Db 135 GCGCAGCGCGAGCTCAGCGATCTGGAGTCGATCTGGCGGGGCAAGAGCTGGCCCGTCG 194

QY 196 LeuProAlaAlaThrProAlaAlaGluLeuValProProAlaHisGluIleGlnGlu 215
 Db 195 CTACCGCGCGCAACCGCGGCTGCGGAGCTGATCTGCGCGCAGCACACGAGATCCAGGAG 254

QY 216 PhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAsp 235
 Db 255 TTCTTCGCGCGCGCGCGAGCGCGCCAGCGCTTTGCTTCCAAAGTACAACTTCGAC 314

QY 236 PheValArgGlyValProLeuAspAlaGlyGlyArgPheGluThrPAlaProValValSer 255
 Db 315 TTGTCGCGCGCGTGCCTTCGACGCGCGCGCGGTTGAGTGGCGCGCGGTGAGC 374

QY 256 Ile 256
 |||

Db 375 ATC 377

RESULT 3
 ABK93954
 ID ABK93954 standard; DNA; 568 BP.

AC ABK93954;

XX 07-AUG-2003 (revised)

DT 27-AUG-2002 (first entry)

XX Corn zmICK1 EST DNA sequence.

DE Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.

XX Zea mays.

XX WO200228893-A2.

XX 11-APR-2002.

XX 29-JUN-2001; 2001WO-IB001492.

XX 14-JUL-2000; 2000US-0218471P.

XX 13-OCT-2000; 2000US-0241219P.

XX (CROP-) CROPDESIGN NV.

XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;

XX Hatzfeld Y;

XX WPI: 2002-471311/50.

XX P-PSDB; ABG65674.

PT Novel plant ICK (inhibitors of Cyclin Dependent Kinases) polypeptide used

PT to screen substrates, drugs or compounds which modulate ICK activity and

PT treat disorders characterized by an insufficient or excessive production

XX of ICK inhibitors.

PS Claim 41; Page 125; 141pp; English.

XX This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a cDNA molecule
 CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
 CC invention. (Updated on 07-AUG-2003 to correct OS field.)

XX SQ Sequence 568 BP; 139 A; 149 C; 177 G; 102 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 6.85e-10 Length: 568
 Score: 21.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.20% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABK93954 (1-568)

QY 225 AlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAla 244
 Db 257 GCCAAGCGCTTCTTCCAAAGTACAACTTCGATTCGCGCGCGGTGCGCGCC 316

QY 245 Gly 245
 |||

Db 317 GGC 319

RESULT 4

AA02390

ID AA02390 standard; DNA; 572 BP.

XX AA02390;

DT 01-NOV-2001 (first entry)

XX Corn Cyclin dependent kinase inhibitor (CDKI) clone cs1ln.pk0050.e6.

DE Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; corn;

XX plant growth inhibitor; ds.

XX Zea mays.

XX Key Location/Qualifiers

FT CDS 208..369

FT /*tag= a

FT /product= "CDKI fragment"

FT /partial

FT /note= "No start or stop codon given"

XX WO200060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TW, Weng Z, Cahoon RE;

XX WPI: 2000-679375/66.

XX P-PSDB; AAP01940.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.
 XX Claim 2; Page 39-40; 58pp; English.
 XX The invention describes a novel isolated polynucleotide comprising a
 CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
 CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
 CC of cell division, growth and death. The nucleotide sequences can be used
 CC in a vector to transform a host cell to produce the CDKI polypeptide.
 CC They can also be used in methods for selecting and obtaining a nucleic
 CC acid sequence that encodes CDKI or affects the level of CDKI expression.
 CC The encoded protein can be used in a method for evaluating a compound for
 CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
 CC as herbicides. They can also be used to inhibit plant growth. The
 CC polynucleotide sequences can be used in gene mapping and as genetic
 CC markers. The sequence encodes the corn CDKI clone cs1n.pk0050.e6 as
 CC described in the method of the invention
 XX SQ Sequence 572 BP; 140 A; 156 C; 176 G; 93 T; 0 U; 7 Other;

Alignment Scores:
 Pred. No.: 6.89e-10 Length: 572
 Score: 21.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.20% Indels: 0
 DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAN02390 (1-572)

QY 225 AlalysArgPheAlaSerIysTrAsnPheAspPheValArgGlyValProleuAspAla 244
 DB 286 GCCAAACGCTTGCTTCCAAAGTACAACTTCGACTTCGTCGGCGGTGCCCTCGACGCC 345

QY 245 Gly 245
 DB 346 GGC 348

RESULT 5
 ID AAA95276 standard; cDNA; 572 BP.

AC AAA95276;

DT 17-JAN-2001 (first entry)

DE Corn cyclin-dependent kinase inhibitor coding sequence #1.

KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 KW cell growth; herbicide; ss.

OS Zea mays.

Key Location/Qualifiers
 CDS 208..366
 FT /*tag= a
 FT /product= "CDKI"
 FT /partial

PN WO200060087-A2.

PD 12-OCT-2000.

PF 06-APR-2000; 2000WO-US009106.

PR 07-APR-1999; 99US-0128192P.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX

DR WPI; 2000-679375/66.
 DR P-PSDB; AAB26245.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.
 XX Claim 2; Page 39-40; 58pp; English.

XX The present sequence is the coding sequence for the corn cyclin-dependent
 CC kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA
 CC library for sequences similar to those encoding the CDKI from Chenopodium
 CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
 CC in the cell cycle, and may promote or inhibit cell division and growth.
 CC The coding sequence and the protein it encodes are useful in the
 CC production of transgenic plants which produce increased or decreased
 CC amounts of the CDKI protein, in the identification of herbicides, in
 CC genetic and physical mapping and in the isolation of the CDKI gene in
 CC other organisms

XX SQ Sequence 572 BP; 140 A; 156 C; 176 G; 93 T; 0 U; 7 Other;

Alignment Scores:
 Pred. No.: 6.89e-10 Length: 572
 Score: 21.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.20% Indels: 0
 DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAA95276 (1-572)

QY 225 AlalysArgPheAlaSerIysTrAsnPheAspPheValArgGlyValProleuAspAla 244
 DB 286 GCCAAACGCTTGCTTCCAAAGTACAACTTCGACTTCGTCGGCGGTGCCCTCGACGCC 345

QY 245 Gly 245
 DB 346 GGC 348

RESULT 6

ID AAA95280 standard; cDNA; 639 BP.

AC AAA95280;

DT 17-JAN-2001 (first entry)

DE Corn cyclin-dependent kinase inhibitor coding sequence #2.

KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 KW cell growth; herbicide; ss.

OS Zea mays.

Key Location/Qualifiers
 CDS 8..388
 FT /*tag= a
 FT /product= "CDKI"
 FT /partial

PN WO200060087-A2.

PD 12-OCT-2000.

PF 06-APR-2000; 2000WO-US009106.

PR 07-APR-1999; 99US-0128192P.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

```
DR P-PSDB; AAB26249.
XX
XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
PT
XX
XX Claim 2; Page 43; 58pp; English.
XX
XX The present sequence is the coding sequence for the corn cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA
CC library for sequences similar to those encoding the CDKI from Chenopodium
CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
CC in the cell cycle, and may promote or inhibit cell division and growth.
CC The coding sequence and the protein it encodes are useful in the
CC production of transgenic plants which produce increased or decreased
CC amounts of the CDKI protein, in the identification of herbicides, in
CC genetic and physical mapping and in the isolation of the CDKI gene in
CC other organisms
XX
XX Sequence 639 BP; 169 A; 169 C; 192 G; 109 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.64e-10 Length: 639
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.20% Indels: 0
DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAA95280 (1-639)
QY 225 AlaLysArgPheAlaSerLysTyrAsnPheValArgGlyValProLeuAspAla 244
Db 293 GCCAAACGCTTGCTTCCAGACACTTCGACTTCGCGGCGTGCCTCGAGGCC 352

QY 245 Gly 245
Db 353 GGC 355

RESULT 7
ABV74605
ID ABV74605 standard; DNA; 841 BP.
XX
XX ABV74605;
AC
XX
DT 21-FEB-2003 (first entry)
DE
DE Maize CKI_D coding sequence.
XX
XX Maize; cyclin-dependent kinase inhibitor; CKI; CKI_D; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication; gene; ds.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
FH 159..839
CDS /*tag= a
FT /partial
FT /product= "CKI_D"
FT /note= "No stop codon given"
XX
XX WO200281623-A2.
XX
XX 17-OCT-2002.
XX
XX 06-NOV-2001; 2001WO-US044038.
XX
XX 07-NOV-2000; 2000US-0246349P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
PI
```

```
XX
XX WPI: 2003-058511/05.
DR P-PSDB; ABB98759.
XX
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX
XX Claim 1; Page 68-69; 69pp; English.
XX
XX The present sequence is the coding sequence (I) for maize cyclin-
CC dependent kinase inhibitor (CKI), CKI_D. (I) is useful for modulating the
CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
CC preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassel size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
CC maize CKI interacting proteins, by aducting the nucleic acid sequence to
CC a second nucleic acid sequence encoding a DNA-binding domain
XX
XX Sequence 841 BP; 142 A; 247 C; 342 G; 109 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 4.8e-06 Length: 841
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.64% Indels: 0
DB: 7 Gaps: 0

US-09-993-808B-2 (1-256) x ABV74605 (1-841)
QY 23 GlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrClyGly 39
Db 225 CAGGTCTCGCGTCCGACGAGGTCCAGGTCCGCGCGCGACCGCGGC 275

RESULT 8
ABK93958
ID ABK93958 standard; DNA; 1242 BP.
XX
XX AC ABK93958;
XX
XX 27-AUG-2002 (first entry)
DT
XX DNA encoding OsICK2 protein.
DE
XX Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
XX
XX Oryza sativa.
OS
XX WO200228893-A2.
XX
XX 11-APR-2002.
XX
XX 29-JUN-2001; 2001WO-IB001492.
XX
XX 14-JUL-2000; 2000US-0218471P.
PR
XX 13-OCT-2000; 2000US-0241219P.
XX
XX (CROP-) CROPDESIGN NV.
XX
```

PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX WPI; 2002-471311/50.
DR P-PSDB; ABG65670.
XX
XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
XX Claim 46; Fig 1; 14lpp; English.
XX
XX This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterized by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterized by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention
XX
SQ Sequence 1242 BP; 286 A; 319 C; 417 G; 220 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.9e-06 Length: 1242
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.64% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABK93958 (1-1242)

Qy 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAla 36
Db 75 GAGGTACAGGAGGTGGTGGCGTCCGACGAGGTGAGTCCGACGGCGG 125

RESULT 9
ABK93981
ID ABK93981 standard; cDNA; 681 BP.
AC ABK93981;
XX
XX 27-AUG-2002 (first entry)
DE Full length Rice OsICK5 cDNA sequence.
XX Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
XX Oryza sativa.
XX
XX WO200228893-A2.
XX
XX 11-APR-2002.
XX
XX 29-JUN-2001; 2001WO-IB001492.
XX
XX 14-JUL-2000; 2000US-0218471P.
XX
XX 13-OCT-2000; 2000US-0241219P.
XX
XX (CROP-) CROPDESIGN NV.
XX
XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX WPI; 2002-471311/50.
DR P-PSDB; ABG65692.
XX
XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
XX Claim 46; Page 140; 14lpp; English.
XX
XX This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterized by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterized by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention
XX
SQ Sequence 681 BP; 102 A; 176 C; 307 G; 94 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 11.1 Length: 681
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABK93981 (1-681)

Qy 237 ValArgGlyValProLeuAspAlaGlyGly 246
Db 613 GTGCGGCGGTGGTGGCTCGACGCGCGCGC 642

RESULT 10
ABK93982
ID ABK93982 standard; DNA; 1073 BP.
AC ABK93982;
XX
XX 27-AUG-2002 (first entry)
DE Rice OsICK5 genomic DNA clone.
XX Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
XX Oryza sativa.
XX
XX WO200228893-A2.
XX
XX 11-APR-2002.
XX
XX 29-JUN-2001; 2001WO-IB001492.
XX
XX 14-JUL-2000; 2000US-0218471P.
XX
XX 13-OCT-2000; 2000US-0241219P.
XX
XX (CROP-) CROPDESIGN NV.
XX
XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX WPI; 2002-471311/50.
XX
XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
XX Claim 46; Page 140; 14lpp; English.
XX
XX This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of

CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a cDNA molecule
 CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
 CC invention

XX SQ Sequence 1073 BP; 172 A; 287 C; 417 G; 197 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 17 Length: 1073
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABX93982 (1-1073)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246

DB 1005 GTGCGGGGGTGGCGTCCGACGCGCGGC 1034

RESULT 11

ACA42627

ID ACA42627 standard; DNA; 1539 BP.

XX AC

XX AC

XX AC

DT 19-JUN-2003 (first entry)

XX AC

DE Prokaryotic essential gene #24284.

XX AC

KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.

XX AC

OS Pseudomonas aeruginosa.

XX AC

PN WO200277183-A2.

XX AC

PD 03-OCT-2002.

XX AC

PF 21-MAR-2002; 2002WO-US009107.

XX AC

PR 21-MAR-2001; 2001US-00815242.

XX AC

PR 06-SEP-2001; 2001US-00948993.

XX AC

PR 25-OCT-2001; 2001US-0342923P.

XX AC

PR 08-FEB-2002; 2002US-00072851.

XX AC

PR 06-MAR-2002; 2002US-0362699P.

XX AC

PA (ELIT-) ELITRA PHARM INC.

XX AC

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX AC

PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;

XX AC

DR WPI; 2003-029926/02.

XX AC

DR P-PSDB; ABU38757.

XX AC

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX AC

PS Claim 14; SEQ ID NO 30497; 1766pp; English.

XX AC

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1539 BP; 197 A; 517 C; 477 G; 349 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 23.8 Length: 1539
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 7 Gaps: 0

US-09-993-808B-2 (1-256) x ACA42627 (1-1539)

QY 113 LeuSerArgCysSerSerThrAlaSerSer 122

DB 998 CTGCTCGTGTTCACAGTACGGCAAGTTCA 1027

RESULT 12

ADA71170

ID ADA71170 standard; DNA; 3525 BP.

XX AC

XX AC

DT 20-NOV-2003 (first entry)

XX AC

DE Rice gene, SEQ ID 4493.

XX AC

KW Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.

XX AC

OS Oryza sativa.

XX AC

PN WO2003000898-A1.

XX AC

PD 03-JAN-2003.

XX AC

PF 22-JUN-2001; 2001WO-IB001105.

XX AC

PR 22-JUN-2001; 2001WO-IB001105.

XX AC

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX AC

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX AC

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX AC

DR WPI; 2003-175290/17.

XX AC

PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX Claim 6; SEQ ID NO 4493; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

SQ Sequence 3525 BP; 777 A; 975 C; 953 G; 814 T; 0 U; 6 Other;

Alignment Scores: 51.6 Length: 3525
Pred. No.: 10.00 Matches: 10
Score: 100.00% Conservations: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 3.91% Gaps: 0
DB: 7

US-09-993-808B-2 (1-256) x ADA71170 (1-3525)

QY 28 ArgThrArgSerArgSerAlaAlaAlaThr 37
DB 206 CGTACGCGCTCGCGCTCTGCGCGCGGACG 235

RESULT 13

AA95286
ID AAA95286 standard; cDNA; 204 BP.

XX AAA95286;

XX 17-JAN-2001 (first entry)

DE Rice cyclin-dependent kinase inhibitor coding sequence #3.

KW Rice; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide; ss.

OS Oryza sativa.

XX Key Location/Qualifiers
FH CDS 11..125
FT /*tag= a
FT /product= "CDKI"
FT /transl_except= (pos:110..113,aa:Ser)
FT /partial

XX WO200060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX P-PSDB; AAB27252.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.

XX

PS Claim 2; Page 48; 58pp; English.

XX The present sequence is the coding sequence for the rice cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a rice seed cDNA
CC library for sequences similar to those encoding the CDKI from Chenopodium
CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
CC in the cell cycle, and may promote or inhibit cell division and growth.
CC The coding sequence and the protein it encodes are useful in the
CC production of transgenic plants which produce increased or decreased
CC amounts of the CDKI protein, in the identification of herbicides, in
CC genetic and physical mapping and in the isolation of the CDKI gene in
CC other organisms

SQ Sequence 204 BP; 42 A; 49 C; 75 G; 29 T; 0 U; 9 Other;

Alignment Scores: 30.3 Length: 204
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservations: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 3.52% Gaps: 0
DB: 3

US-09-993-808B-2 (1-256) x AAA95286 (1-204)

QY 237 ValArgGlyValProLeuAspAlaGly 245
DB 59 GTTCGCGCGTGGCCCTCGACGCGGT 85

RESULT 14

AA952400
ID AA952400 standard; DNA; 204 BP.

XX AA952400;

XX 01-NOV-2001 (first entry)

DE Rice Cyclin dependent kinase inhibitor (CDKI) clone rds2c.px008.o24.

KW Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; rice;
KW plant growth inhibitor; ds.

OS Oryza sativa.

XX Key Location/Qualifiers
FH CDS 11..125
FT /*tag= a
FT /product= "CDKI fragment"
FT /partial
FT /note= "No start codon given"
FT /transl_except= (pos:107..110,aa:Val)
FT /note= "This codon has an apparent 1 nucleotide insertion
FT which alters the reading frame"

XX WO200060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX P-PSDB; AAP01950.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.

XX Claim 2; Page 48; 58pp; English.

XX The invention describes a novel isolated polynucleotide comprising a
 CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
 CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
 CC of cell division, growth and death. The nucleotide sequences can be used
 CC in a vector to transform a host cell to produce the CDKI polypeptide.
 CC They can also be used in methods for selecting and obtaining a nucleic
 CC acid sequence that encodes CDKI or affects the level of CDKI expression.
 CC The encoded protein can be used in a method for evaluating a compound for
 CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
 CC as herbicides. They can also be used to inhibit plant growth. The
 CC polynucleotide sequences can be used in gene mapping and as genetic
 CC markers. The sequence encodes the rice CDKI clone rds2c.pk008.o24 as
 CC described in the method of the invention

XX SQ Sequence 204 BP; 42 A; 49 C; 75 G; 29 T; 0 U; 9 Other;

Alignment Scores:
 Pred. No.: 30.3 Length: 204
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.52% Indels: 0
 DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAN02400 (1-204)

QY 237 ValArgGlyValProLeuAspAlaGly 245

DB 59 GTTCGGCGGTGCGCCCTCGACGCGGT 85

RESULT 15

ABL81995

ID ABL81995 standard; cDNA; 334 BP.

XX ABL81995;

XX 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:4973.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide.

XX Claim 1; SEQ ID NO 4973; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
 CC or antigen presenting cells that express (II). (I) has cytostatic
 CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for

CC detecting ovarian cancer in a patient's biological sample preferably
 CC serum or ovarian tissue. The method comprises contacting a biological
 CC sample from a patient with (IV), detecting the amount of polynucleotide
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff
 CC value and thereby detecting ovarian cancer in the patient, where the
 CC amount of polynucleotide hybridising to (IV) is detected preferably by
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
 CC useful for stimulating and/or expanding T cells specific for an ovarian
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is
 CC useful in design and preparation of ribozyme molecules for inhibiting
 CC expression of the tumour polypeptides and proteins in tumour cells; and
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
 CC library using well known techniques

XX SQ Sequence 334 BP; 61 A; 87 C; 145 G; 41 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 47.9 Length: 334
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.52% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABL81995 (1-334)

QY 10 GlyAlaAlaGlyAlaGluValAlaAla 18

DB 159 GGAGCGCGCGTCTGAGGTGGCGGCC 185

Search completed: October 2, 2004, 13:10:26

Job time : 373 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 12:56:31 ; Search time 85 Seconds
(without alignments)
1671.382 Million cell updates/sec

Title: US-09-993-808B-2

Perfect score: 256

Sequence: 1 MGKYMCKRGAAGAAVE.....VRGVLDAAGREWAPVWSI 256

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360453

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0993808/runat_01102004_171727_13734/app_query.fasta_1.455
-DB=Issued Patents NA -QFW=fastap -SUFFIX=oligo.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808 @CNG 1.1 105 @runat_01102004_171727_13734 -NCPH=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=6 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	3.5	436	US-09-302-769-30	Sequence 30, Appl
2	9	3.5	642	US-09-252-991A-6820	Sequence 6820, Ap
3	9	3.5	714	US-09-252-991A-6929	Sequence 6929, Ap
4	9	3.5	723	US-09-252-991A-4826	Sequence 4826, Ap
5	9	3.5	771	US-09-489-039A-4276	Sequence 4276, Ap
6	9	3.5	792	US-09-489-039A-6808	Sequence 6808, Ap
7	9	3.5	1053	US-09-252-991A-6360	Sequence 6360, Ap
8	9	3.5	1548	US-09-252-991A-4907	Sequence 4907, Ap
9	9	3.5	1602	US-09-252-991A-4836	Sequence 4836, Ap
10	9	3.5	1623	US-09-252-991A-6190	Sequence 6190, Ap
11	9	3.5	2052	US-09-252-991A-6112	Sequence 6112, Ap
12	9	3.5	34279	US-09-596-002-26	Sequence 26, Appl

c	13	8	3.1	372	4	US-09-134-000C-2276	Sequence 2276, Ap
	14	8	3.1	378	4	US-09-235-996B-3	Sequence 3, Appli
	15	8	3.1	378	4	US-09-295-846B-3	Sequence 3, Appli
	16	8	3.1	378	4	US-09-551-737C-3	Sequence 3, Appli
	17	8	3.1	378	4	US-09-551-738B-3	Sequence 3, Appli
	18	8	3.1	422	3	US-09-060-756-568	Sequence 568, App
	19	8	3.1	422	4	US-09-670-314-568	Sequence 568, App
	20	8	3.1	447	4	US-09-252-991A-10231	Sequence 10231, A
	21	8	3.1	471	2	US-08-883-070-2	Sequence 2, Appli
	22	8	3.1	471	4	US-09-621-976-847	Sequence 847, App
	23	8	3.1	483	4	US-09-252-991A-7028	Sequence 7028, Ap
	24	8	3.1	520	4	US-09-621-976-3031	Sequence 3031, Ap
	25	8	3.1	540	4	US-09-252-991A-8395	Sequence 8395, Ap
	26	8	3.1	654	4	US-09-252-991A-14070	Sequence 14070, A
	27	8	3.1	683	4	US-09-885-166A-896	Sequence 896, App
	28	8	3.1	705	1	US-10-095-946-17	Sequence 17, Appl
	29	8	3.1	705	4	US-09-183-959-17	Sequence 17, Appl
	30	8	3.1	705	4	US-09-535-315-17	Sequence 17, Appl
	31	8	3.1	750	1	US-10-095-946-18	Sequence 18, Appl
	32	8	3.1	750	4	US-09-183-959-18	Sequence 18, Appl
	33	8	3.1	750	4	US-09-535-315-18	Sequence 18, Appl
	34	8	3.1	764	4	US-08-426-630-21	Sequence 21, Appl
	35	8	3.1	765	4	US-09-252-991A-1872	Sequence 1872, Ap
	36	8	3.1	846	4	US-09-252-991A-8327	Sequence 8327, Ap
	37	8	3.1	1001	4	US-09-016-434-963	Sequence 963, App
	38	8	3.1	1017	4	US-09-252-991A-10799	Sequence 10799, A
	39	8	3.1	1040	1	US-10-095-946-11	Sequence 11, Appl
	40	8	3.1	1040	4	US-09-183-959-11	Sequence 11, Appl
	41	8	3.1	1040	4	US-09-535-315-11	Sequence 11, Appl
	42	8	3.1	1077	4	US-09-252-991A-10719	Sequence 10719, A
	43	8	3.1	1101	4	US-09-252-991A-10349	Sequence 10349, A
	44	8	3.1	1113	4	US-09-252-991A-8445	Sequence 8445, Ap
	45	8	3.1	1173	4	US-09-252-991A-10448	Sequence 10448, A

ALIGNMENTS

RESULT 1
US-09-302-769-30
Sequence 30, Application US/09302769
Patent No. 6323317
GENERAL INFORMATION:
APPLICANT: HILTON, Douglas J
APPLICANT: ALEXANDER, Warren S
APPLICANT: VINEY, Elizabeth M
APPLICANT: WILLSON, Tracey A
APPLICANT: RICHARDSON, Rachael T
APPLICANT: STARR, Robyn
APPLICANT: NICHOLSON, Sandra E
APPLICANT: METCALF, Donald
APPLICANT: NICOLA, Nicos A
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
FILE REFERENCE: 10976Z
CURRENT APPLICATION NUMBER: US/09/302,769
CURRENT FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 08/962,560
PRIOR FILING DATE: 1997-10-31
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 30
LENGTH: 436
TYPE: DNA
ORGANISM: Human
US-09-302-769-30

Alignment Scores:
Pred. No.: 25.4 Length: 436
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.52% Indels: 0
DB: Gaps: 0

Db 150 GCGCTGTGAGAGACACGAGCTGCAGCT 176

RESULT 6

US-09-489-039A-6808

; Sequence 6808, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 6808

; LENGTH: 792

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-6808

Alignment Scores:

Pred. No.:	43.4	Length:	792
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.52%	Indels:	0
DB:	4	Gaps:	0

US-09-993-808B-2 (1-256) x US-09-489-039A-6808 (1-792)

QY 50 AlaProAlaGlyGluProAlaAlaA 58

|||||

Db 150 GCGCTGTGAGAGACACGAGCTGCAGCT 176

RESULT 7

US-09-252-991A-6360

; Sequence 6360, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6360

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6360

Alignment Scores:

Pred. No.:	56	Length:	1053
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.52%	Indels:	0
DB:	4	Gaps:	0

US-09-993-808B-2 (1-256) x US-09-252-991A-6360 (1-1053)

QY 153 valSerAlaAlaSerAsnSerGlySer 161

|||||

Db 413 GTGTGGCGCGGTGGAACAGCGGTTC 439

RESULT 8

US-09-252-991A-4907/c

; Sequence 4907, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4907

; LENGTH: 1548

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4907

Alignment Scores:

Pred. No.:	79	Length:	1548
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.52%	Indels:	0
DB:	4	Gaps:	0

US-09-993-808B-2 (1-256) x US-09-252-991A-4907 (1-1548)

QY 114 SerArgCysSerSerThrAlaSerSer 122

|||||

Db 563 TCTCGCTGTTCAGTACGCGAAGTTCA 537

RESULT 9

US-09-252-991A-4836

; Sequence 4836, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4836

; LENGTH: 1602

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4836

Alignment Scores:

Pred. No.:	81.5	Length:	1602
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.52%	Indels:	0
DB:	4	Gaps:	0

US-09-993-808B-2 (1-256) x US-09-252-991A-4836 (1-1602)

QY 114 SerArgCysSerSerThrAlaSerSer 122

|||||

Db 1064 TCTCGCTGTTCAGTACGCGAAGTTCA 1090

RESULT 10

US-09-252-991A-6190/c

```
; Sequence 6190, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6190
; LENGTH: 1623
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6190
Alignment Scores:
Pred. No.: 82.4 Length: 1623
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.52% Indels: 0
DB: 4 Gaps: 0
US-09-993-808B-2 (1-256) x US-09-252-991A-6190 (1-1623)
Qy 153 ValSerAlaAlaSerAsnSerGlySer 161
Db 1244 GTGTCGGCGGCGTCGAACAGCGGTTC 1218
RESULT 11
US-09-252-991A-6112/c
; Sequence 6112, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6112
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6112
Alignment Scores:
Pred. No.: 102 Length: 2052
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.52% Indels: 0
DB: 4 Gaps: 0
US-09-993-808B-2 (1-256) x US-09-252-991A-6112 (1-2052)
Qy 153 ValSerAlaAlaSerAsnSerGlySer 161
Db 1578 GTGTCGGCGGCGTCGAACAGCGGTTC 1552
RESULT 12
US-09-596-002-26/c
; Sequence 26, Application US/09596002
```

```
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 60/140,121
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 34279
; TYPE: DNA
; ORGANISM: M. catarrhalis
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 26
; PUBLICATION INFORMATION:
US-09-596-002-26
Alignment Scores:
Pred. No.: 1.26e+03 Length: 34279
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.52% Indels: 0
DB: 4 Gaps: 0
US-09-993-808B-2 (1-256) x US-09-596-002-26 (1-34279)
Qy 229 AlaSerLysTyrAsnPheAspPheVal 237
Db 15679 GCGTCAATATATATTTGATTGTA 15653
RESULT 13
US-09-134-000C-2276/c
; Sequence 2276, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2276
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2276
Alignment Scores:
Pred. No.: 164 Length: 372
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.12% Indels: 0
DB: 4 Gaps: 0
US-09-993-808B-2 (1-256) x US-09-134-000C-2276 (1-372)
Qy 115 ArgCysSerSerThrAlaSerSer 122
Db 107 CGATGCTCTTCCACCGCATCTTC 84
RESULT 14
```



```
US-09-295-996B-3
; Sequence 3, Application US/09295996B
; Patent No. 6413530
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: PESTICIDAL PEPTIDES
; FILE REFERENCE: UF-230
; CURRENT APPLICATION NUMBER: US/09/295,996B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Aedes aegypti
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(189)
; OTHER INFORMATION:
US-09-295-996B-3

Alignment Scores:
Pred. No.: 166 Length: 378
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.12% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-295-996B-3 (1-378)
QY 154 SerAlaAlaSerAsnSerGlySer 161
Db 112 AGTGTGCTCCCAACTCGGGTGC 135

RESULT 15
US-09-295-846B-3
; Sequence 3, Application US/09295846B
; Patent No. 6562590
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: Transformed Cells Useful for the Control of Pests
; FILE REFERENCE: UF-223
; CURRENT APPLICATION NUMBER: US/09/295,846B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Aedes aegypti
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(189)
US-09-295-846B-3

Alignment Scores:
Pred. No.: 166 Length: 378
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.12% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-295-846B-3 (1-378)
QY 154 SerAlaAlaSerAsnSerGlySer 161
Db 112 AGTGTGCTCCCAACTCGGGTGC 135

Search completed: October 2, 2004, 14:45:07
Job time : 96 secs
```

This Page Blank (uspto)

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 2, 2004, 13:10:36 ; Search time 446 Seconds
        (without alignments)
        2909.886 Million cell updates/sec

Title:
Perfect score: 256
Sequence: 1 MGKYMKRCRGAAGAEVAAVE.....VRCVPLDAGRPFAVPVSI 256

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3340653 seqs, 2534783454 residues
Word size: 1
Total number of hits satisfying chosen parameters: 6672345

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0993808/runat_01102004_171728_13817/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0993808@cgn_1_1_723@runat_01102004_171728_13817
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LOGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10E_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US10K_PUBCOMB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US10L_PUBCOMB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US10M_PUBCOMB.seq:*
27: /cgn2_6/ptodata/2/pubpna/US10N_PUBCOMB.seq:*
28: /cgn2_6/ptodata/2/pubpna/US10O_PUBCOMB.seq:*
29: /cgn2_6/ptodata/2/pubpna/US10P_PUBCOMB.seq:*
30: /cgn2_6/ptodata/2/pubpna/US10Q_PUBCOMB.seq:*
31: /cgn2_6/ptodata/2/pubpna/US10R_PUBCOMB.seq:*
32: /cgn2_6/ptodata/2/pubpna/US10S_PUBCOMB.seq:*
33: /cgn2_6/ptodata/2/pubpna/US10T_PUBCOMB.seq:*
34: /cgn2_6/ptodata/2/pubpna/US10U_PUBCOMB.seq:*
35: /cgn2_6/ptodata/2/pubpna/US10V_PUBCOMB.seq:*
36: /cgn2_6/ptodata/2/pubpna/US10W_PUBCOMB.seq:*
37: /cgn2_6/ptodata/2/pubpna/US10X_PUBCOMB.seq:*
38: /cgn2_6/ptodata/2/pubpna/US10Y_PUBCOMB.seq:*
39: /cgn2_6/ptodata/2/pubpna/US10Z_PUBCOMB.seq:*
40: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
41: /cgn2_6/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
42: /cgn2_6/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
43: /cgn2_6/ptodata/2/pubpna/US11D_PUBCOMB.seq:*
44: /cgn2_6/ptodata/2/pubpna/US11E_PUBCOMB.seq:*
45: /cgn2_6/ptodata/2/pubpna/US11F_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
------------	-------------	--------	-------	-------------

ALIGNMENTS

```
RESULT 1
US-09-993-308-1
; Sequence 1, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (134)...(904)
; OTHER INFORMATION:
US-09-993-308-1

Alignment Scores:
Pred. No.: 1,49e-233 Length: 1372
Score: 256.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-993-308-1 (1-1372)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
DB 134 ATGGGAAAGTACATCGCAAGTGCAGGCGGCCCGCAGCGCGAGGTGCGCCGCGTCGAG 193

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
DB 194 GTTACGCGAGGTCTGCGGCTCCGACGAGGTCCAGTCCGCGCGCGGCGGCGGTGTC 253

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
DB 254 CGCAAGGTTCGCCCGCGAGGAGAGGCGCGCGCGGCGGAGCCTGCTGCGCGCGTGAGC 313

QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
DB 314 GCTGTGGGGACGCGGGAAGTGTACATCCACCTGCTAGCGCGCATCTGTTCATGGCA 373

QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
DB 374 CCGCTTCAGCGCGACCGCTCGGTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGGC 433

QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
DB 434 GCTGCGGACACGAGGCGCGCGCTCGCGTTCGACCCCGGTGGAGGCTGCTGATGGC 493

QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
DB 494 TCGTCGGTGAACCTTGGGCTTGGGGGTTCAGCGCGGAGCCACCTCGCGCTCCCTACGAC 553

QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
DB 554 GCTGCGGACGCGCGGAGGCGAGACGACGATCCAGAGTTCCTGCGCGCGCGC 613

QY 161 SerGlyProAspAlaGlyArgPheGluThrProAlaProValSerIle 256
DB 794 GAGCGGCGCGGCGGCGGCTTGGTTCGAGTGGGCGCGCGTGGTGGTTCGAGCATC 853

QY 241 ProLeuAspAlaGlyGlyArgPheGluThrProAlaProValSerIle 256
DB 854 CCCCTCGACGCGGCGCGGCTTGGTTCGAGTGGGCGCGCGTGGTGGTTCGAGCATC 901
```

RESULT 2

```
US-09-993-808B-1
; Sequence 1, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
```

```
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)...(902)
US-09-993-808B-1

Alignment Scores:
Pred. No.: 1,49e-233 Length: 1372
Score: 256.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
```

```
US-09-993-808B-2 (1-256) x US-09-993-808B-1 (1-1372)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
DB 134 ATGGGAAAGTACATCGCAAGTGCAGGCGGCCCGCAGCGCGAGGTGCGCCGCGTCGAG 193

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
DB 194 GTTACGCGAGGTCTGCGGCTCCGACGAGGTCCAGTCCGCGCGCGGCGGCGGTGTC 253

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
DB 254 CGCAAGGTTCGCCCGCGAGGAGAGGCGCGCGCGGCGGAGCCTGCTGCGCGCGTGAGC 313

QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
DB 314 GCTGTGGGGACGCGGGAAGTGTACATCCACCTGCTAGCGCGCATCTGTTCATGGCA 373

QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
DB 374 CCGCTTCAGCGCGACCGCTCGGTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGGC 433

QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
DB 434 GCTGCGGACACGAGGCGCGCGCTCGCGTTCGACCCCGGTGGAGGCTGCTGATGGC 493

QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
DB 494 TCGTCGGTGAACCTTGGGCTTGGGGGTTCAGCGCGGAGCCACCTCGCGCTCCCTACGAC 553

QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
DB 554 GCTGCGGACGCGCGGAGGCGAGACGACGATCCAGAGTTCCTGCGCGCGCGC 613

QY 161 SerGlyProAspAlaGlyArgPheGluThrProAlaProValSerIle 180
DB 614 AGCGGCGCGAGCGGAGGAGAGACGACGATCCAGCGCGCGCGGCGGAGGCTC 673

QY 181 SerAspLeuGluSerAspLeuAlaGlyHisTyrThrGlyProSerLeuProAlaAlaThr 200
DB 674 AGCGATCTGGAGTCCGATCTCGCGGCGGCGACCAAGACTGGCCCGCTCGTACCACGCAACG 733

QY 201 ProAlaAlaGluLeuValProAlaHisGluIleGlnGluPhePheAlaAlaAla 220
DB 734 CCGGCTCGGAGGCTGATCTGCTGCGCGGCGGCGGAGGATCCAGGAGTTCCTGCGCGCGCGG 793
```

```
QY 221 GluAlaAlaGluAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
Db 794 GAGGCGGCGCCAGCCCAAGCGCTTTGTTCCAAAGTACAACCTTCGACTTCGTCGCGCGGTG 853
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrrPAlaProValValSerIle 256
Db 854 CCCTTCGACCGCGCGCGGTTTCGAGTGGCGCGCGTGTGTGTCAGCATC 901

RESULT 3
US-10-425-114-18379/c
; Sequence 18379, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18379
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI
US-10-425-114-18379

Alignment Scores:
Pred. No.: 1,316-101 Length: 985
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 45.70% Indels: 0
DB: 13 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-425-114-18379 (1-985)

QY 140 AspAlaAlaGluAlaGlyGlyArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 159
Db 672 GACGCTGCGAGGCTGGCGGGGATCACGCTCTGTGTGATGTCTCGCGCGCGCAACTCC 613
QY 160 GlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGlu 179
Db 612 GGGAGCGGCGCCAGACCGCGAGAGCGGAGACGACGCCATCGAGCGCGGCGCGCGGAG 553
QY 180 LeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAla 199
Db 552 CTCAGCGATCTGAGTCGAGTCTGGCGGGGCAACAAGACTGGCCCGTCTACCGCGGCA 493
QY 200 ThrProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPheAlaAla 219
Db 492 ACCCGCGCTCGAGGTGATCGTGGCGCCAGCACAGAGATCCAGAGATTCCTTCGCGGCC 433
QY 220 AlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGly 239
Db 432 GCGGAGCGGCGCCAGCGCCGCTTTGCTTCCAAAGTACAACCTTCGACTTCGTCGCGTGGC 373
QY 240 ValProLeuAspAlaGlyGlyArgPheGluTrrPAlaProValValSerIle 256
Db 372 GTGCCCTTCAGCGCGCGCGGTTTCGAGTGGCGCGCGTGTGTGTCAGCATC 322

RESULT 4
US-10-425-114-34947
; Sequence 34947, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18379
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI
US-10-425-114-18379
```

```
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34947
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17355F05_FLI
US-10-425-114-34947

Alignment Scores:
Pred. No.: 1,456-101 Length: 1111
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 45.70% Indels: 0
DB: 13 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-425-114-34947 (1-1111)

QY 140 AspAlaAlaGluAlaGlyGlyArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 159
Db 440 GACGCTGCGAGGCTGGCGGGGATCACGCTCTGTGTGATGTCTCGCGCGCGCAACTCC 499
QY 160 GlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGlu 179
Db 500 GGGAGCGGCGCCAGACCGCGAGAGCGGAGACGACGCCATCGAGCGCGGCGCGGAG 559
QY 180 LeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAla 199
Db 560 CTCAGCGATCTGAGTCGAGTCTGGCGGGGCAACAAGACTGGCCCGTCTACCGCGGCA 619
QY 200 ThrProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPheAlaAla 219
Db 620 ACCCGCGCTCGAGGTGATCGTGGCGCCAGCACAGAGATCCAGAGATTCCTTCGCGGCC 679
QY 220 AlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGly 239
Db 680 GCGGAGCGGCGCCAGCGCCGCTTTGCTTCCAAAGTACAACCTTCGACTTCGTCGCGTGGC 739
QY 240 ValProLeuAspAlaGlyGlyArgPheGluTrrPAlaProValValSerIle 256
Db 740 GTGCCCTTCAGCGCGCGCGGTTTCGAGTGGCGCGCGTGTGTGTCAGCATC 790

RESULT 5
US-10-767-701-24179
; Sequence 24179, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24179
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
```

RESULT 7

```
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-333-006-9

Alignment Scores:
Pred. No.: 1.61e-06      Length: 1242
Score: 17.00           Matches: 17
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 6.64%              Indels: 0
DB: 16                    Gaps: 0

US-09-993-808B-2 (1-256) x US-10-333-006-9 (1-1242)

QY 20 GluValThrGlnValValGlyValAlaThrArgSerArgSerAlaAla 36
Db 75 GAGGTCCAGCAGGTGGTGGCGTCCGACGAGGTCCGACGCGCGC 125

RESULT 11
US-10-437-963-92632/c
; Sequence 92632, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 92632
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1
US-10-437-963-92632

Alignment Scores:
Pred. No.: 2.05e-06      Length: 1671
Score: 17.00           Matches: 17
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 6.64%              Indels: 0
DB: 17                    Gaps: 0

US-09-993-808B-2 (1-256) x US-10-437-963-92632 (1-1671)

QY 20 GluValThrGlnValValGlyValAlaThrArgSerArgSerAlaAla 36
Db 1149 GAGGTCCAGCAGGTGGTGGCGTCCGACGAGGTCCGACGCGCGC 1099

RESULT 12
US-10-027-632-149457/c
; Sequence 149457, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
US-09-993-808B-2 (1-256) x US-09-993-308-5 (1-841)

QY 23 GlnValValGlyValAlaThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db 225 CAGGTGTCGGCGTCGGACGAGGTCCAGGTCCGCGCGCGCGCGCGC 275

RESULT 9
US-09-993-808B-5
; Sequence 5, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 841
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(841)
; OTHER INFORMATION: r= g or a
US-09-993-808B-5

Alignment Scores:
Pred. No.: 1.18e-06      Length: 841
Score: 17.00           Matches: 17
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 6.64%              Indels: 0
DB: 11                    Gaps: 0

US-09-993-808B-2 (1-256) x US-09-993-808B-5 (1-841)

QY 23 GlnValValGlyValAlaThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db 225 CAGGTGTCGGCGTCGGACGAGGTCCAGGTCCGCGCGCGCGCGC 275

RESULT 10
US-10-333-006-9
; Sequence 9, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1242
```

US-09-993-808B-2 (1-256) x US-10-027-632-149457 (1-803)
QY 107 AlalalaLeuAlaAlaGlyLeuSerArgCysSer 117
DB 135 GCAGCGTGGCGGAGGACTGAGCGGTGCTCC 103
RESULT 14
US-10-333-006-54
; Sequence 54, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 54
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: probe or primer
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(356)
; OTHER INFORMATION: n = A, T, C or G
US-10-333-006-54

Alignment Scores:
Pred. No.: 0.568 Length: 803
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.30% Indels: 0
DB: 13 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-027-632-149457 (1-803)

QY 107 AlalalaLeuAlaAlaGlyLeuSerArgCysSer 117
DB 135 GCAGCGTGGCGGAGGACTGAGCGGTGCTCC 103

RESULT 13

US-10-027-632-149457/c
; Sequence 149457, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149457
; LENGTH: 803
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-149457

Alignment Scores:
Pred. No.: 0.568 Length: 803
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.30% Indels: 0
DB: 16 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-027-632-149457 (1-803)
QY 107 AlalalaLeuAlaAlaGlyLeuSerArgCysSer 117
DB 135 GCAGCGTGGCGGAGGACTGAGCGGTGCTCC 103
RESULT 14
US-10-333-006-54
; Sequence 54, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 54
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: probe or primer
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(356)
; OTHER INFORMATION: n = A, T, C or G
US-10-333-006-54

Alignment Scores:
Pred. No.: 4.43 Length: 681
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 16 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-333-006-54 (1-681)
QY 237 ValargglyValProLeuAspAlaGlyGly 246
DB 613 GTGCGCGCGGTGCGCGTGTGCGCGCGGCG 642
RESULT 15
US-10-437-963-19949
; Sequence 19949, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

Alignment Scores:
Pred. No.: 4.43 Length: 681
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 16 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-333-006-54 (1-681)
QY 237 ValargglyValProLeuAspAlaGlyGly 246
DB 613 GTGCGCGCGGTGCGCGTGTGCGCGCGGCG 642
RESULT 15
US-10-437-963-19949
; Sequence 19949, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

Alignment Scores:
Pred. No.: 4.43 Length: 681
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 16 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-333-006-54 (1-681)
QY 237 ValargglyValProLeuAspAlaGlyGly 246
DB 613 GTGCGCGCGGTGCGCGTGTGCGCGCGGCG 642
RESULT 15
US-10-437-963-19949
; Sequence 19949, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963


```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 19949
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2535C.1
US-10-437-963-19949
```

```
Alignment Scores:
Pred. No.: 6.37 Length: 1065
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 17 Gaps: 0
```

US-09-993-808B-2 (1-256) x US-10-437-963-19949 (1-1065)

```
QY 237 ValArgGlyValProLeuAspAlaGlyGly 246
Db 682 GTGCGGGCGTGGCGCTCGACGCGCGGCG 711
```

Search completed: October 2, 2004, 15:47:27
Job time : 451 secs

This Page Blank (uspto)